Statistical Analysis Plan:

11-gene Correlates of Risk (COR) Diagnostic and Predictive Performance Analysis in HIV-Infected Adults

Protocol Title:	Validation of Correlates of Risk of TB Disease in High Risk Populations (CORTIS-HR) *A companion study of the CORTIS-01 Trial*
Protocol Number:	CORTIS-HR
Authors:	Simon C Mendelsohn[1], MBChB Andrew Fiore-Gartland[2], Ph.D. Bhavesh Borate[2], M.S. [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:	11 December 2019
Version:	v 1.0

This document describes the primary statistical analysis plan to be performed on the data obtained from the CORTIS-HR study and is to be read in conjunction with the approved study protocol version 1.0 dated 26th August 2016.

Document: Statistical Analysis Plan: 11-gene Correlates of Risk (COR) Diagnostic and Predictive Performance Analysis in HIV-Infected Adults. Final version 1.0 (Dated 11 December 2019)

Protocol Title: Validation of Correlates of Risk of TB Disease in High Risk Populations (CORTIS-HR)

A companion study of the CORTIS-01 Trial

Protocol Number: CORTIS-HR

Authorised by: Prof Mark Hatherill

Director, South African Tuberculosis Vaccine Initiative, University of Cape Town

Sign: 

Date: 12 DEC 2019

Dr Andrew Fiore-Gartland

Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center

Sign: 

Date: Dec 11-2019

Table of Contents

1 List of Abbreviations ____ *5*

2 Study Overview ____ *6*

Aims ____ 6

2.1 Study Design ____ **7**

2.2 Study Population ____ **7**

3 Statistical Considerations ____ *8*

3.1 Statistical software ____ **8**

3.2 Responsibility ____ **8**

3.3 Blinding ____ **8**

3.4 Database lock and storage ____ **8**

3.5 Projected enrolment, case accrual and power ____ **8**

3.6 COR predictive performance analysis ____ **10**

3.7 TB disease endpoint adjudication algorithm ____ **10**

Raw endpoint data ____ 11

Study endpoints, censoring and cohort definitions ____ 12

3.8 Statistical methods ____ **14**

3.8.1 Methods for primary, secondary, and exploratory aims ____ 14

3.8.1 Classification of participants by endpoint/censoring status ____ 16

3.9 Tables and figures ____ **17**

3.9.1 Descriptive analyses ____ 17

3.9.2 COR and QFT performance analyses ____ 17

4 Tables ____ *19*

4.1 Participant disposition (consort diagram) ____ **19**

4.2 Endpoint accrual ____ **20**

4.3 Enrolment and endpoint accrual by COR, ART and IPT status ____ **21**

4.4 Enrolled participant demographics and clinical characteristics, by site ____ **22**

4.5 Enrolled participant demographics and clinical characteristics By case status with crude odds ratio ____ **26**

4.6 Primary and secondary outcomes tables (BINARY ANALYSIS) ____ **31**

4.6.1 Cumulative test performance for primary end-point (two-sample) prevalent and incident TB cases during follow-up (Primary aim) ____ 31

4.6.2 Cumulative test performance for combined primary and secondary end-point (one- or two-sample) prevalent and incident TB cases during follow-up (Primary aim) 31
4.6.3 Screening/triage test performance for primary end-point (two-sample) prevalent TB cases at enrolment (Secondary aim 1) 31
4.6.4 Screening/triage test performance for combined primary and secondary end-point (one- or two-sample) prevalent TB cases at enrolment (Secondary aim 1) 31
4.6.5 Predictive test performance for primary end-point (two-sample) incident TB cases during follow-up 31
4.6.6 Predictive test performance for combined primary and secondary end-point (one- or two-sample) incident TB cases during follow-up 31

4.7 Performance of the COR (RISK11) and QFT for identification of TB disease over a 15-month period, stratified by the time interval to disease (Time-dependent analysis) 34
4.7.1 Cumulative test performance for primary end-point (two-sample) prevalent and incident TB cases during follow-up (Primary aim) 34
4.7.2 Cumulative test performance for combined primary and secondary end-point (one- or two-sample) prevalent and incident TB cases during follow-up (Primary aim) 34
4.7.3 Predictive test performance for primary end-point (two-sample) incident TB cases during follow-up (Secondary aims 2 and 3) 34
4.7.4 Predictive test performance for combined primary and secondary end-point (one- or two-sample) incident TB cases during follow-up (Secondary aims 2 and 3) 34

5 Figures 35

5.1 Types of Figures 35
5.1.1 Test performance: ROC curves (sensitivity versus 100-specificity) 35
5.1.2 Test accuracy: sensitivity/specificity versus test score threshold plots (x-axis: score threshold, y-axis: sensitivity or specificity) 35
5.1.3 Test score distribution: violin/box-and-scatter plots (Prevalent and incident TB cases versus TB negative controls) 35
5.1.4 Time dependent analysis: AUC versus time. CIR versus time. Sensitivity, specificity, NPV and PPV (at different thresholds) versus time, over 15-month follow-up 35

5.2 Endpoints for figures 35

5.3 Tests 35
5.3.1 COR (RISK11) 35
5.3.2 QuantiFERON 35

6 Appendix 36

6.1 WHO diagnostic, triage, and predictive TB test performance target product profile 36

6.2 Concomitant medication coding 37

7 References 43

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 STUDY OVERVIEW

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 with equivalent diagnostic performance (Darboe, *Tuberculosis*, 2018).

COR diagnostic performance for discriminating prevalent TB disease from latent TB infection in HIV infected persons appears to be reduced by approximately 10%, compared to HIV uninfected persons (Darboe, *Front Microbiol*, 2019). The overarching objective of this study is to evaluate the performance of the 11-gene mRNA COR signature to identify prevalent TB disease and predict incident TB disease in HIV infected individuals.

AIMS

Primary aim

Test whether COR status differentiates HIV infected persons with cumulative prevalent or incident TB disease from those without TB disease.

Secondary aims

1. Estimate whether COR status differentiates HIV infected persons with prevalent TB disease from those without prevalent TB disease
2. Estimate whether COR status differentiates HIV infected persons at high risk for incident TB disease from those at low risk for incident TB disease
3. Compare predictive performance of the COR for incident TB disease with Interferon-gamma release assay (IGRA) in HIV infected persons.

Exploratory aims

(Not addressed in this statistical analysis plan)

1. Assess and model the impact of a COR screen & treat strategy on reducing the rate of incident TB disease and TB mortality among HIV infected persons in South Africa.
2. Re-parameterize the COR assay for prevalent and incident TB disease in HIV infected persons.
3. Test the performance of additional validated COR signatures in distinguishing HIV infected persons with cumulative prevalent or incident TB disease from those without TB disease.
4. Compare predictive performance of the COR for incident TB disease in HIV infected persons before and 3 months after starting IPT and/or ART.

2.1 STUDY DESIGN

CORTIS-HR is an observational study of COR diagnostic and predictive performance in HIV infected persons. Following sample collection for the COR assay at Visit 1 (Day 0), all participants will be assessed for prevalent TB with TB investigations. All eligible participants without prevalent TB will be referred for Isoniazid preventive therapy (IPT). All HIV infected persons eligible for ART, as per SA national guidelines, will be referred in writing to the clinic for ART. Thereafter, all participants will be followed for 15 months for incident TB disease (3 telephone contacts and 4 site visits). Symptoms consistent with TB disease will be solicited at study visits and presence of one or more symptoms will trigger TB investigation. TB investigations will also be performed on all participants at the end of study follow up (month 15). The performance of the COR will be evaluated by comparing the cumulative incidence of endpoint-defined TB disease over 15 months in COR+ versus COR- participants (RR_{COR}).

2.2 STUDY POPULATION

The study population will include 860 HIV infected adults residing in TB hyperendemic communities at five or more study sites in South Africa. Primarily, persons screened for the CORTIS-01 trial who are ineligible on the grounds of HIV infection will be approached by study staff to participate in CORTIS-HR. In addition, persons with documented HIV infection may be approached through local ART clinics and Voluntary Counselling & Testing (VCT) services affiliated with or run directly by the study sites. Although recruitment efforts will be focused on persons with documented HIV infection status, in order to maintain confidentiality and avoid risk of stigmatization, community volunteers with unknown HIV status would also be eligible for screening.

3 STATISTICAL CONSIDERATIONS

3.1 STATISTICAL SOFTWARE

Statistical analysis will be performed using statistical software including Prism (GraphPad Software Inc.) and R. Where significance is reported, the 5% level of significance will be used with correction for false discovery and multiple comparisons.

3.2 RESPONSIBILITY

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

3.3 BLINDING

The researcher assigned to study oversight and statistical analysis will be blinded to participant COR scores, COR status, and QuantiFERON (QFT) results until this statistical analysis plan has been approved and signed by study sponsor. However, access to clinical and demographic data (include TB disease status) will available to allow data cleaning and preparation of analysis script prior to database lock.

3.4 DATABASE LOCK AND STORAGE

After completion of study follow up, acquisition of final sputum culture results, and completion of database cleaning by Triclinium Clinical Development (TCD), database will be locked by agreement of the study PI, sponsor, and analysis team. All files containing the final data will be password-protected and backed-up to the SATVI server under a “CORTIS-HR” folder. 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.

3.5 PROJECTED ENROLMENT, CASE ACCRUAL AND POWER

We estimated the expected number of prevalent and incident TB cases in HIV infected participants in CORTIS-HR, based on data from previous and ongoing studies in South Africa. In order to estimate the expected number of TB cases and a probable range, we used stochastic simulations of the CORTIS-HR study under a range of assumptions. The simulation assumed a 10% HIV prevalence rate, ranging between 5% - 25%, which would allow 860 HIV infected participants to be enrolled from approximately 10,000 persons screened for the CORTIS trial and the CORTIS-HR study.

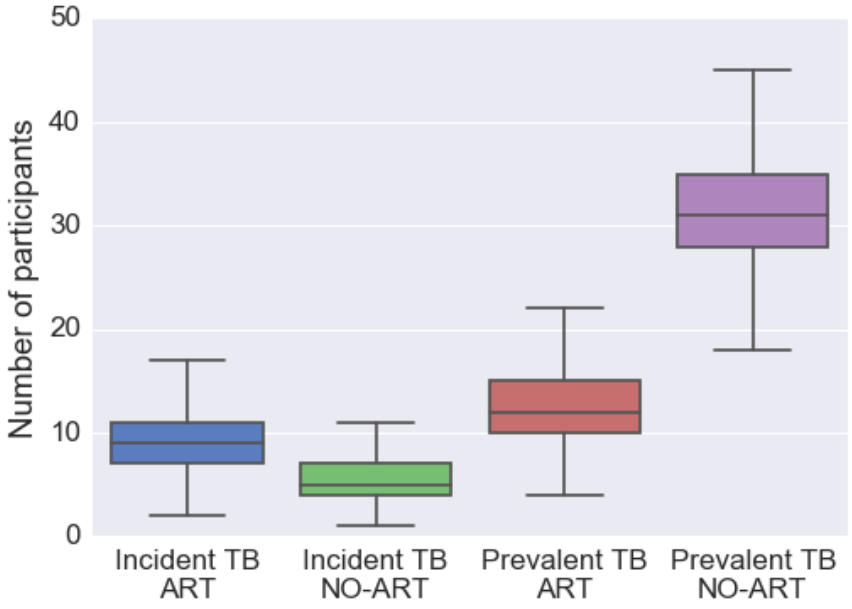


Figure 1: Expected number of incident and prevalent TB endpoints among HIV infected participants. Estimates are based on stochastic simulations.

We estimate that 75% of these participants will be receiving chronic ART and the remainder will begin ART. Based on studies conducted in South Africa, the rate of prevalent TB will be considerably higher among participants newly starting ART (15%), compared to those receiving chronic ART (2%). We also assume that 15% of participants on chronic ART will also be on IPT, thus further reducing the expected level of prevalent TB by 40%. Similarly, we expect the annual rate of incident TB will be higher among participants newly starting ART (4%), than among those already established on chronic ART (2%). In both these groups, we expect up to 70% to begin IPT upon enrolment, which will reduce expected levels of incident TB by 40%.

Based on these pre-specified parameter distributions, we expect that of the 215 newly diagnosed HIV infected participants yet to start ART, we will identify 31 prevalent TB cases (95% CI 22, 42) at screening; and an additional 5 (95% CI 2, 10) participants who will develop active TB disease over 15 months of follow-up. We expect that of the 645 HIV infected participants already on chronic ART, we will identify 12 prevalent TB cases (95% CI 6, 20) at screening; and an additional 9 (95% CI 4, 15) incident TB cases over 15 months of follow-up. Thus, we expect to measure the COR in a total of 58 (95% CI 44, 73) cumulative endpoint TB cases (43 prevalent and 14 incident TB cases) and 802 (95% CI 787, 816) controls.

3.6 COR PREDICTIVE PERFORMANCE ANALYSIS

In the CORTIS-01 study, a pre-defined COR score threshold of 60% was used to differentiate correlate positive (COR+) from correlate negative (COR-) individuals. There was no pre-defined COR threshold for HIV infected individuals, thus we plan to test multiple COR thresholds (See **Table 4.6**). For Secondary Aim 3, we will also assess multiple thresholds for the IGRA (See **Table 4.6**). To assess COR predictive performance for the prediction of incident TB disease we will evaluate the relative-risk of endpoint-defined TB over 15 months, $RR_{COR}(15)$, in COR+ versus COR- participants using a cumulative incidence-based approach. The primary analysis will evaluate $RR_{COR}(15)$ on endpoints using the two-sample endpoint definition and according to the formula:

$$RR_{COR}(15) = \frac{H_A}{H_B}$$

where H_X is the cumulative incidence estimated for each group using the Product-Limit estimator (Aalen, *The Annals of Statistics*, 1978). A point-estimate for $RR_{COR}(15)$ will be presented with 95% confidence intervals and a p-value for the null-hypothesis H_0: $RR_{COR}(15) \leq 1$. In addition we will present time-dependent estimates of sensitivity, specificity, positive predictive value (PPV) and number needed to treat (NNT) using the methods of Heagerty (*Biometrics*, 2000) as these will offer important insights into the performance and application of the COR in future strategies.

3.7 TB DISEASE ENDPOINT ADJUDICATION ALGORITHM

The algorithm is needed to classify each sputum sample as positive or negative and subsequently, each participant as TB-negative, two-sample positive (primary TB disease endpoint) or one-sample positive (secondary TB disease endpoint). It also establishes the study visit and timepoint at which TB-positive participants will be considered TB-positive.

The algorithm makes explicit two basic rules of adjudication: (1) Thirty-day episode window, and (2) “worst, first” case definition. The 30-day episode window implements the concept that assay results from multiple samples should be considered related and possibly combined to indicate a two-sample positive endpoint if they occur within 30-days of the first positive sample within the episode; samples collected more than 30 days apart should be considered as independent episodes for endpoint adjudication. The “first, worst” rule refers to the concept that a participant should be classified as a two-sample TB-positive endpoint if at any point during the study there is a two-sample positive episode, even if it follows a previous one-sample positive episode.

A window of 14 days will be applied to the end-of-study month 15 visit (i.e. a sputum collected ≥ month 15 + 14 days will be excluded).

RAW ENDPOINT DATA

Each participant in the study provides a number of samples and associated assay results. The visit number and collection dates are indicated for each sample by the BARC_MICRO.visit and BARC_MICRO.coll_dt, respectively. We consider two assays and their associated positive results below. GeneXpert MTB/RIF and Xpert Ultra “Trace” results are excluded from primary analysis.

Assay	Value of BARC_MICRO.proc	Value of BARC_MICRO.prompt	Value of BARC_MICRO.res indicating positive result
Mycobacteria Growth Indicator Tube (MGIT)	MGIT 960 Mycobacterial Culture	M tuberculosis :	Positive
	MGIT 960 Mycobact Re-Culture	M tuberculosis :	Positive
GeneXpert MTB/RIF	MTB PCR GENE EXPERT	Organism 1	Mycobact tuberculosis complex.
	MTB PCR GENE EXPERT ULTRA	Organism 1	Mycobact tuberculosis complex.
	MTB PCR GENE EXPERT ULTRA	Organism 1	Mycobact tuberculosis complexu

Apply the following steps to each participant’s set of assay results, up until the time of the analysis:

Step 1. If there are no positive assay results the participant is classified as TB negative for all Study Visits. Proceed to adjudicate the next participant.

Step 2. Begin with the first Study Visit with an associated sample that is positive for either assay. This becomes the Episode Start Visit and starting date.

Step 3. If there are at least two samples collected at the Episode Start Visit that are positive for TB by either assay, then classify the participant as two-sample TB-positive at that visit. Proceed to adjudicate the next participant.

Step 4. If there is only one positive sample from the Episode Start Visit, examine all samples within a 30-day period. If there are any subsequent samples that are TB positive by either assay, then classify the participant as two-sample TB-positive at the Episode Start Visit. Proceed to adjudicate the next participant.

Step 5. If the participant has no subsequent TB-positive results, then classify the participant as one-sample TB-positive at the Episode Start Visit. Proceed to adjudicate the next participant.

Step 6. If the participant has TB-positive samples remaining, establish a new Episode Start Visit at the next positive sample (this should correspond to the next planned Study Visit). Continue with Step 3 to consider this new Episode.

STUDY ENDPOINTS, CENSORING AND COHORT DEFINITIONS

A participant classified as a one-sample or two-sample positive TB endpoint at the Enrolment Visit (Visit 1) will be classified as a prevalent TB case. Participants classified as a one-sample or two-sample positive TB endpoint at a subsequent visit will be defined as incident TB disease cases with right-censoring at the date of the positive sputum sample collection. Individuals who remain TB negative until study discontinuation, or end of follow up at the month 15 (day 449) end of study visit are classified as TB negative controls.

The primary and secondary aims will be addressed through time dependent and binary analyses of the intention to treat (ITT) cohort and modified ITT (mITT) cohorts:

The ITT cohort will include all participants, with censoring at discontinuation visit (e.g. unscheduled visit for pregnancy) or at last real visit (a real visit is one where a TB symptom screen was conducted or BARC microbiological data is available). For two-sample (primary endpoint), only consider two-sample endpoints as endpoints. One-sample endpoints will be censored at their last real visit. The ITT cohort will be used for addressing the primary aim (COR performance in differentiating cumulative prevalent and incident cases from TB negative controls) and secondary aim 1 (COR diagnostic performance in differentiating prevalent cases from combined TB negative controls and incident TB cases).

The mITT cohort excludes prevalent TB cases and will be used for assessing the performance of COR and QFT in predicting progression to TB disease (secondary aims 2 and 3).

Primary Analysis	Cohort	Cases	Controls
Primary aim: Cumulative prevalent and incident cases (time-dependent endpoints)	ITT	Incident and prevalent TB	TB negative
Secondary aim 1: Diagnostic performance (binary endpoints)	ITT	Prevalent TB	Incident TB and TB negative
Secondary aims 2 and 3: Predictive performance (time-dependent endpoints)	mITT (excluding prevalent cases)	Incident TB	TB negative

An exploratory analysis using binary endpoints will also be conducted and will only include participants who met a TB endpoint (prevalent or incident) or attended the month 15 end of study (EOS) visit without early discontinuation or left censoring.

Exploratory Analysis	Cohort	Cases	Controls
Diagnostic performance (binary endpoints)	Exclude early discontinuation and incident TB from ITT	Prevalent TB	TB negative
Predictive performance (binary endpoints)	Exclude early discontinuation and prevalent TB from ITT	Incident TB	TB negative

3.8 STATISTICAL METHODS

3.8.1 METHODS FOR PRIMARY, SECONDARY, AND EXPLORATORY AIMS

<table>
<tr><td rowspan="2">Endpoint*</td><td rowspan="2">Statistic</td><td colspan="3">Relative-risk (COR+ vs. COR-, IGRA+ vs. IGRA-)</td><td colspan="4">ROC statistics (COR, IGRA)</td></tr>
<tr><td>Cumulative ITT</td><td>Predictive (incident) mITT (exclude V1 TB)</td><td>Diagnostic ITT</td><td>Sensitivity</td><td>Specificity</td><td>AUC</td><td>Others (e.g. NNS, PPV, NPV)</td></tr>
<tr><td rowspan="3">Primary analysis: Time-dependent right censored</td><td>Statistic</td><td colspan="2">CIR</td><td>-</td><td colspan="4">survAM.estimate function, survAccuracyMeasures package (Zheng, 2016)
estimation.method= IPW (non parametric, inverse-probability weighted)</td></tr>
<tr><td>95% CI</td><td colspan="2">Wald-based</td><td>-</td><td colspan="4">Bootstrapping, 10,000 replicates, non-stratified, non-parametric resampling</td></tr>
<tr><td>Test (p-value)</td><td colspan="2">Wald-based
H_0: $CIR_{COR}(15) = 1$</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td></tr>
<tr><td rowspan="3">Exploratory analysis: Binary#</td><td>Statistic</td><td colspan="3">RR</td><td colspan="4">pROC package (Robin, BMC Bioinformatics, 2011)</td></tr>
<tr><td>95% CI</td><td colspan="3">Likelihood-score based (Koopman, Biometrics, 1984; Nam, Biometrical Journal, 1995)
PropCIs package (Scherer, 2018)</td><td colspan="4">Percentile method (Carpenter, Statistics in Medicine, 2000),
10,000 bootstrap replicates with non-stratified, non-parametric resampling</td></tr>
<tr><td>Test (p-value)</td><td colspan="3">Chi squared,
H_0: $RR_{COR}(15) = 1$</td><td>Comparing sensitivity/specificity of two tests:
(Pepe, Stata J, 2009), roc.test function, pROC package,
H_0: $Sensitivity_{COR} = Sensitivity_{IGRA}$
H_0: $Specificity_{COR} = Specificity_{IGRA}$</td><td colspan="2">Single AUC: Mann-Whitney U (wilcox.test),
H_0: AUC < 0.5

Comparing AUCs: roc.test function (DeLong, Biometrics, 1988), pROC package,
H_0: $AUC_{COR} = AUC_{IGRA}$</td><td>-</td></tr>
</table>

*All analyses will be conducted twice: (I) Two-sample endpoints, (II) One and two-sample endpoints combined.

[#] In circumstances when there are 0 cases in one or more critical pieces of a binary endpoint evaluation (i.e. true positives or false negatives), 1 person will be added to each cell of the 2x2 table to help stabilise the estimate. This is particularly pertinent to the exploratory sub-group analyses: 1) ART-naïve versus ART experienced at baseline, 2) CD4 cell count at baseline <500, versus ≥ 500, 3) viral load at baseline lower than the detectable limit (LDL, <100), versus ≥ 100, 4) prior TB episode versus no prior TB episode, and 5) those who received IPT during study, versus no IPT during study. Other exploratory subgroups analyses will include age (<35, ≥35), sex (male/female), ethnicity (Black African, Cape Mixed Ancestry), study site, smoking history (yes/no), BMI (<18.5, ≥18.5), baseline IPT status (yes/no), and presence of TB symptoms (yes/no). For the RISK11 (COR) signature, an additional sub-group analysis of baseline QFT level will be performed (<0.35, ≥0.35; and <4, ≥4). These analyses may have limited power due to small sample size and limited numbers of active TB cases.

3.8.1 CLASSIFICATION OF PARTICIPANTS BY ENDPOINT/CENSORING STATUS

No. of samples	Endpoint?	V2 ENDP?	$T_{endpoint}/T_{LRV}$?	Assign T_{event}	Assign *endpoint*	Label
One (OS) or Two (TS) sample	Endpoint	≥V2	$T_{endpoint} \geq 15.5^{*}$	T_{LRV-1} (LRV‡ prior to endpoint)	0	OS_ENDP_>V2_T>15.5
			$14.5 \geq T_{endpoint} < 15.5$	15	1	OS_ENDP_>V2_T15
			$T_{endpoint} < 14.5$	$T_{endpoint}$	1	OS_ENDP_>V2_T<14.5
		V1	-			OS_ENDP_V1
	Non-endpoint	-	$T_{LRV} \geq 14.5$	15	0	OS_NON_FULL
			$T_{LRV} < 14.5$	T_{LRV}	0	OS_NON
Two sample only	Same classification#, but only consider Two-sample endpoints as endpoints. One-sample endpoints will be censored at their LRV.					TS_ENDP_>V2_T>15.5
						TS_ENDP_>V2_T15
						TS_ENDP_>V2_T<14.5
						TS_ENDP_V1
						TS_NON_FULL
						TS_NON

‡LRV = last real visit, may include EOSM15 only if participant COMPLETED the study and was present at the EOSM15 visit, evidenced by presence of vital signs, TB screening, or BARC microbiological data

$T_{endpoint}$ = date determined from the endpoint adjudication process, specific to the endpoint (i.e. one- or two-sample)

T_{event} = elapsed time to endpoint or censor, for use in downstream analyses

*T > 15.5: a two-week visit window will be enforced for final visits that occur more than two weeks past 15 months. Also LRVs/endpoints occurring within two weeks before or after 15 months will be rounded to precisely 15 months to ensure a stable risk pool at the primary analysis timepoint

#Classification is valid if all ITT participants can be classified using one category for OS+TS endpoint and one category for TS endpoint

3.9 TABLES AND FIGURES

3.9.1 DESCRIPTIVE ANALYSES

Total number of subjects consented, screened, passing inclusion criteria, tested for COR, and enrolled as well as not enrolled into the trial will be summarized by site and COR status. Trial completion, trial withdrawals, exclusions and protocol non-compliances will also be summarized (**Table 4.1**).

Primary and secondary incident and prevalent endpoint accrual will be summarized by site and by COR status for subjects in the ITT cohort (**Table 4.2**).

Endpoint accrual based on two- and one-sample detection will also be summarized by ART and IPT treatment, and COR status (**Table 4.3**). ART and IPT status is determined by the concomitant medication recorded during follow up (**Appendix 6.2**).

Enrolled participant demographic, clinical and laboratory characteristics will be summarized by site (**Table 4.4**). Pearson's Chi-squared test (categorical data) or Kruskal-Wallis (continuous data) will be used to test for any statistically significant differences in demographic or baseline data between sites. A p value of <0.05 will be considered statistically significant.

Enrolled participant demographic, clinical and laboratory characteristics will also be summarized by case status (prevalent, incident and cumulative TB cases, and non-TB-endpoint controls) in the ITT cohort (**Tables 4.5**). Chi-squared test (categorical data) or logistic regression (continuous data) will be used to test for any statistically significant differences in demographic or baseline data between cases and controls. A p value of <0.05 will be considered statistically significant.

3.9.2 COR AND QFT PERFORMANCE ANALYSES

The performance of the COR and QFT (**primary and secondary aims**), 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 using the following performance measures: area under the receiver operating characteristic curve (AUC), relative-risk of COR+ vs. COR- (RR_{COR}), sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and the number needed to screen to detect one case (NNS).

COR performance will be compared with symptom screening, QFT, and urinary lipoarabinomannan (LAM) lateral flow assays also measured at screening. The estimates and confidence intervals will be provided in **Tables 4.6.1-6** alongside the minimum and optimal Target Product Profiles (TPP), provided by the WHO, for a community-based triage or referral test to identify people suspected of having TB (World

Health Organization, 2014, 2017). The WHO diagnostic, triage, and predictive TB test performance target product profiles (TPP) are included as **Appendix 6.1**.

Assessment of cumulative (primary aim), diagnostic (secondary aim 1), and predictive (secondary aim 2) performance of COR and QFT (secondary aim 3) will be based on only two-sample detection endpoints as well as both two- and one-sample detection endpoints (**Tables 4.6.1-6**). Different COR and QFT thresholds will be evaluated. ROC curves (**Figure 5.1.1**), sensitivity/specificity versus test score threshold plots (**Figure 5.1.2**), and violin/box-and-scatter plots (**Figure 5.1.3**) will be presented to describe COR and QFT performance at different cut-offs and to demonstrate score distributions.

Longitudinal performance of the COR and QFT assays (secondary aims 2 and 3) for identification of TB disease over a 15-month period based on only two-sample detection endpoints as well as both two- and one-sample detection endpoints, will be stratified by the time interval to disease (**Table 4.7.1-4** and **Figure 5.1.4**). The time dependent longitudinal analysis will allow estimation of AUC versus time, as well as sensitivity, specificity, NPV, and PPV (at a-priori and post-hoc thresholds) versus time, over 15-month follow-up. The optimal threshold will be set based on 75% test specificity (minimum WHO TPP for a predictive test for progression to TB disease).

Pre-specified exploratory diagnostic and predictive performance analyses (**Tables 4.6.1-6**, **Tables 4.7.1-4** and **Figures 5.1.1-5.1.4**) will also be performed by the following sub-groups: 1) ART-naïve versus ART experienced at baseline, 2) CD4 cell count at baseline <500, versus ≥ 500, 3) viral load at baseline lower than the detectable limit (LDL, <100), versus ≥ 100, 4) prior TB episode versus no prior TB episode, and 5) those who received IPT during study, versus no IPT during study. Other exploratory subgroups analyses will include age (<35, ≥35), sex (male/female), ethnicity (Black African, Cape Mixed Ancestry), study site, smoking history (yes/no), BMI (<18.5, ≥18.5), baseline IPT status (yes/no), and presence of TB symptoms (yes/no). For the RISK11 (COR) signature, an additional sub-group analysis of baseline QFT level will be performed (<0.35, ≥0.35; and <4, ≥4). These analyses may have limited power due to small sample size and limited numbers of active TB cases.

4 TABLES

4.1 PARTICIPANT DISPOSITION (CONSORT DIAGRAM)

Category	Subcategory		AIKlerks	AIRusten	CAPRISA	SATVI	SUN	Total
Consented			x	x	x	x	x	x
Not screened	*Out of enrolment window or did not arrive for enrolment visit*		x	x	x	x	x	x
Screened		N	x	x	x	x	x	x
Excluded		n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	Aged <18 or ≥60	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	HIV negative	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	Pregnant or lactating	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	Unlikely to remain in follow up and adhere to protocol requirements	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	Diagnosed with TB disease within last 3 years	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	Household exposure to known MDR-TB patient within last 3 years	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	Any medical, surgical, or other condition likely to interfere with mRNA signature performance, safety or efficacy of ART and/or IPT, or adherence to study protocol	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	Missing inclusion data	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Enrolled		N	x	x	x	x	x	x
	COR+	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	COR-	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	COR indeterminate	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	COR missing	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
End of study	All scheduled visits attended	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	Missed visits but no early termination	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	Total early discontinuation / termination	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	Prevalent TB case (diagnosed in study)	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	Incident TB case (diagnosed in study)	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	TB case (diagnosed and treated for TB externally)	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	Consent withdrawal	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	Withdrawal by investigator	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	LTFU	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	Death	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	Pregnancy	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	Protocol deviation/violation	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	Other	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	Missing end of study reason	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)

Note: Percentages for Category column are computed using total screened or total enrolled as denominator. Percentages for Subcategory column are computed using total in the particular Category column as denominator.

4.2 ENDPOINT ACCRUAL

		AIKlerks		AIRusten		CAPRISA		SATVI		SUN		Total	
		COR+	COR-	COR+	COR-	COR+	COR-	COR+	COR-	COR+	COR-	COR+	COR-
Total subjects enrolled	N	x	x	x	x	x	x	x	x	x	x	x	x
Prevalent TB: Subjects diagnosed with TB at enrolment/visit 1*	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
1. Primary endpoint: based on two-sample detection	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Positive Xpert MTB/RIF	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Positive MGIT culture	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Positive Xpert MTB/RIF & MGIT culture	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
2. Secondary endpoint: based on one-sample detection	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Positive Xpert MTB/RIF	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Positive MGIT culture	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Positive Xpert MTB/RIF & MGIT culture	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Incident TB: Subjects diagnosed with TB at Visits 2-8*	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
1. Primary endpoint: based on two-sample detection	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Positive Xpert MTB/RIF	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Positive MGIT culture	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Positive Xpert MTB/RIF & MGIT culture	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
2. Secondary endpoint: based on one-sample detection	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Positive Xpert MTB/RIF	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Positive MGIT culture	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Positive Xpert MTB/RIF & MGIT culture	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Enrolled participants with TB diagnosed on clinical grounds, or not having a one- or two-sputum sample positive endpoint (excluded from diagnostic analysis)	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)

Note: * Percentages in this group are computed using total subjects enrolled (Row 1) as denominator. All subjects with one-sample detection had more than one specimen tested.

4.3 ENROLMENT AND ENDPOINT ACCRUAL BY COR, ART AND IPT STATUS

		ART+			ART-			Total		
		IPT+	**IPT-**	**Total**	**IPT+**	**IPT-**	**Total**	**IPT+**	**IPT-**	**Total**
Total subjects enrolled	N	x	x	x	x	x	x	x	x	x
COR+	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
COR-	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
COR indeterminate*	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Subjects diagnosed with prevalent TB at enrolment/visit 1	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Based on two-sample detection	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Based on one-sample detection	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Subjects diagnosed with incident TB during follow-up	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Based on two-sample detection	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Based on one-sample detection	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
COR+	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Subjects diagnosed with prevalent TB at enrolment/visit 1	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Based on two-sample detection	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Based on one-sample detection	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Subjects diagnosed with incident TB during follow-up	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Based on two-sample detection	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Based on one-sample detection	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
COR-	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Subjects diagnosed with prevalent TB at enrolment/visit 1	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Based on two-sample detection	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Based on one-sample detection	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Subjects diagnosed with incident TB during follow-up	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Based on two-sample detection	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Based on one-sample detection	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
COR indeterminate	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Subjects diagnosed with prevalent TB at enrolment/visit 1	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Based on two-sample detection	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Based on one-sample detection	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Subjects diagnosed with incident TB during follow-up	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Based on two-sample detection	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Based on one-sample detection	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)

Note: Percentages are computed using total number in enrolled cohort as denominator.

4.4 ENROLLED PARTICIPANT DEMOGRAPHICS AND CLINICAL CHARACTERISTICS, BY SITE

Variable	Levels		Total	AIKlerks	AIRusten	CAPRISA	SATVI	SUN	p value
Total enrolled		N	x	x	x	x	x	x	
Demographics									
Gender	Female	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	Male	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
Age		median (IQR)	x (x–x)	x (x–x)	x (x–x)	x (x–x)	x (x–x)	x (x–x)	p
Ethnicity	Black	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	Mixed Race	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
Highest education level	Primary School or Lower	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	Secondary School	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
	Tertiary	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
Employment status	Employed	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	Unemployed	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
Occupants per household		median (IQR)	x (x–x)	x (x–x)	x (x–x)	x (x–x)	x (x–x)	x (x–x)	p
Occupants per room		median (IQR)	x (x–x)	x (x–x)	x (x–x)	x (x–x)	x (x–x)	x (x–x)	p
Smoking history	Never	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	Current	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	Past	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
History of prior TB	No	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	Yes	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
TB household contacts	No	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	Yes	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
Season of enrolment	Summer	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	Autumn	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
	Winter	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
	Spring	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)	

Clinical variables								
Weight (kg), baseline		median (IQR)	x (x–x)	x (x–x)	x (x–x)	x (x–x)	x (x–x)	p
	Missing data	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
Weight (kg), month 3		median (IQR)	x (x–x)	x (x–x)	x (x–x)	x (x–x)	x (x–x)	p
	Missing data	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
Body-mass index, baseline		median (IQR)	x (x–x)	x (x–x)	x (x–x)	x (x–x)	x (x–x)	p
	Missing data	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	<18.5	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	18.5-24.9	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
	>25	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
Weight change: M3-D0		median (IQR)	x (x–x)	x (x–x)	x (x–x)	x (x–x)	x (x–x)	p
Weight change: EOS-D0		median (IQR)	x (x–x)	x (x–x)	x (x–x)	x (x–x)	x (x–x)	p
Antiretroviral therapy (ART) at baseline	Started After Enrolment	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	<6 Months	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
	6-12 Months	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
	>12 Months	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
	No ART Recorded	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
Isoniazid Preventive Therapy (IPT) at baseline	Started After Enrolment	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	<6 Months	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
	6-12 Months	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
	>12 Months	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
	No IPT Recorded	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
Laboratory results								
QuantiFERON, baseline		median (IQR)	x (x–x)	x (x–x)	x (x–x)	x (x–x)	x (x–x)	p
	QuantiFERON not done/missing	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	QuantiFERON Indeterminate	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	Negative (<0.35)	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	Low positive (0.35 to<1.0)	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
	Medium positive (1.0 to<4.0)	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
	High positive (≥4.0)	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	

QuantiFERON, baseline (binary)	Negative (<0.35)	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	Positive (≥0.35)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	
CD4 cell count, baseline		median (IQR)	x (x–x)	x (x–x)	x (x–x)	x (x–x)	x (x–x)	p
	CD4 cell count not done	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	<100	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	100-349	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
	350-499	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
	≥500	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
CD4 cell count, baseline (binary)	<500	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	≥500	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
CD4 cell count, month 3		median (IQR)	x (x–x)	x (x–x)	x (x–x)	x (x–x)	x (x–x)	p
	CD4 not done/missing	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	<100	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	100-349	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
	350-499	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
	≥500	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
CD4 cell count, month 3 (binary)	<500	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	≥500	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
Viral load, baseline	<100	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	100-999	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
	≥1000	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
	Viral load not done/missing	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
Viral load, month 3	<100	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	100-999	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
	≥1000	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
	Viral load not done/missing	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
COR score, baseline		median (IQR)	x (x–x)	x (x–x)	x (x–x)	x (x–x)	x (x–x)	p
	Indeterminate	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	Not done/missing	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
COR score, month 3		median (IQR)	x (x–x)	x (x–x)	x (x–x)	x (x–x)	x (x–x)	p
	Indeterminate	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p

	Not done/missing	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
Urinary LAM, baseline	Negative	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	Positive	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
	Not done/missing	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
Urinary LAM, month 3	Negative	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	Positive	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
	Not done/missing	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
TB Symptoms								
Cough > 2 Weeks	No	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	Yes	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
Haemoptysis in past 2 weeks	No	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	Yes	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
Weight loss > 2 weeks	No	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	Yes	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
Fever > 2 weeks	No	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	Yes	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
Night sweats > 2 weeks	No	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	Yes	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
Chest pain > 2 weeks	No	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	Yes	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
Flu-like symptoms in past 2 weeks	No	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	Yes	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
	Not recorded	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
WHO symptom screen	Negative	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	Positive	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
Number of TB symptoms	0	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	p
	1	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
	2	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
	3	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	
	4	n (%)	x (%)	x (%)	x (%)	x (%)	x (%)	

4.5 ENROLLED PARTICIPANT DEMOGRAPHICS AND CLINICAL CHARACTERISTICS BY CASE STATUS WITH CRUDE ODDS RATIO

Variable	Levels		Healthy	Cumulative TB cases	OR (95% CI)	p value	Prevalent TB cases	OR (95% CI)	p value	Incident TB cases	OR (95% CI)	p value
Total enrolled		N	x	x			x			x		
Demographics												
Gender	Female	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	Male	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
Age		median (IQR)	x (x–x)	x (x–x)	x (x–x) per 10 years	p	x (x–x)	x (x–x) per 10 years	p	x (x–x)	x (x–x) per 10 years	p
Ethnicity	Black	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	Mixed Race	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
City/Site	Worcester, Western Cape	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	Durban, KwaZulu Natal	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
	Klerksdorp, North West	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
	Rustenburg, North West	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
	Stellenbosch, Western Cape	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
Highest education level	Primary School or Lower	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	Secondary School	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
	Tertiary	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
Employment status	Employed	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	Unemployed	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
Occupants per household		median (IQR)	x (x–x)	x (x–x)	x (x–x) per occupant	p	x (x–x)	x (x–x) per occupant	p	x (x–x)	x (x–x) per occupant	p
Occupants per room		median (IQR)	x (x–x)	x (x–x)	x (x–x) per occupant	p	x (x–x)	x (x–x) per occupant	p	x (x–x)	x (x–x) per occupant	p
Smoking history	Never	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	Current	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
	Past	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
History of prior TB	No	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	Yes	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
TB household contacts	No	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	

	Yes	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
Season of enrolment	Summer	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	Autumn	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
	Winter	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
	Spring	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
Clinical variables												
Weight (kg), baseline		median (IQR)	x (x–x)	x (x–x)	x (x–x) per 10kg	p	x (x–x)	x (x–x) per 10kg	p	x (x–x)	x (x–x) per 10kg	p
	Missing data	n (%)	x (%)	x (%)			x (%)			x (%)		
Weight (kg), month 3		median (IQR)	x (x–x)	x (x–x)	x (x–x) per 10kg	p	x (x–x)	x (x–x) per 10kg	p	x (x–x)	x (x–x) per 10kg	p
	Missing data	n (%)	x (%)	x (%)			x (%)			n (%)		
Body-mass index, baseline		median (IQR)	x (x–x)	x (x–x)	x (x–x) per unit	p	x (x–x)	x (x–x) per unit	p	x (x–x)	x (x–x) per unit	p
	Missing data	n (%)	x (%)	x (%)			x (%)			x (%)		
	<18.5	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	18.5-24.9	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
	>25	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
Weight change: M3-D0		median (IQR)	x (x–x)	x (x–x)	x (x–x) per kg	p	x (x–x)	x (x–x) per kg	p	x (x–x)	x (x–x) per kg	p
Weight change: EOS-D0		median (IQR)	x (x–x)	x (x–x)	x (x–x) per kg	p	x (x–x)	x (x–x) per kg	p	x (x–x)	x (x–x) per kg	p
Antiretroviral therapy (ART) at baseline	Started After Enrolment	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	<6 Months	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
	6-12 Months	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
	>12 Months	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
	No ART Recorded	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
Isoniazid Preventive Therapy (IPT) at baseline	Started After Enrolment	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	<6 Months	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
	6-12 Months	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
	>12 Months	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
	No IPT Recorded	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p

Laboratory results												
QuantiFERON, baseline		median (IQR)	x (x–x)	x (x–x)	x (x–x) per 50 cells	p	x (x–x)	x (x–x) per 50 cells	p	x (x–x)	x (x–x) per 50 cells	p
	QuantiFERON done/missing	n (%)	x (%)	x (%)			x (%)			x (%)		
	QuantiFERON indeterminate	n (%)	x (%)	x (%)			x (%)			x (%)		
	Negative (<0.35)	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	Low positive (0.35 to<1.0)	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
	Medium positive (1.0 to<4.0)	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
	High positive (≥4.0)	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
QuantiFERON, baseline (binary)	Negative (<0.35)	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	Positive (≥0.35)	n (%)	n (%)	n (%)	x (x–x)	p	n (%)	x (x–x)	p	n (%)	x (x–x)	p
CD4 cell count, baseline		median (IQR)	x (x–x)	x (x–x)	x (x–x) per 50 cells	p	x (x–x)	x (x–x) per 50 cells	p	x (x–x)	x (x–x) per 50 cells	p
	CD4 not done/missing	n (%)	x (%)	x (%)			x (%)			x (%)		
	<100	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	100-349	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
	350-499	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
	≥500	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
CD4 cell count, baseline (binary)	<500	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	≥500	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
CD4 cell count, month 3		median (IQR)	x (x–x)	x (x–x)	x (x–x) per 50 cells	p	x (x–x)	x (x–x) per 50 cells	p	x (x–x)	x (x–x) per 50 cells	p
	CD4 not done/missing	n (%)	x (%)	x (%)			x (%)			x (%)		
	<100	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	100-349	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
	350-499	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
	≥500	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
CD4 cell count, month 3 (binary)	<500	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	≥500	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
Viral load, baseline	<100	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	100-999	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
	≥1000	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p

	Viral load not done/missing	n (%)	x (%)	x (%)			x (%)			x (%)		
Viral load, month 3	<100	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	100-999	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
	≥1000	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
	Viral load not done/missing	n (%)	x (%)	x (%)			x (%)			x (%)		
COR score, baseline		median (IQR)	x (x–x)	x (x–x)	x (x–x) per 10%	p	x (x–x)	x (x–x) per 10%	p	x (x–x)	x (x–x) per 10%	p
	Indeterminate	n (%)	x (%)	x (%)			x (%)			x (%)		
	Not done/missing	n (%)	x (%)	x (%)			x (%)			x (%)		
COR score, month 3		median (IQR)	x (x–x)	x (x–x)	x (x–x) per 10%	p	x (x–x)	x (x–x) per 10%	p	x (x–x)	x (x–x) per 10%	p
	Indeterminate	n (%)	x (%)	x (%)			x (%)			x (%)		
	Not done/missing	n (%)	x (%)	x (%)			x (%)			x (%)		
Urinary LAM, baseline	Negative	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	Positive	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
	Not done/missing	n (%)	x (%)	x (%)			x (%)			x (%)		
Urinary LAM, month 3	Negative	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	Positive	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
	Not done/missing	n (%)	x (%)	x (%)			x (%)			x (%)		
Symptom screening												
Cough > 2 Weeks	No	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	Yes	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
Haemoptysis in past 2 weeks	No	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	Yes	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
Weight loss > 2 weeks	No	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	Yes	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
Fever > 2 weeks	No	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	Yes	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
Night sweats > 2 weeks	No	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	Yes	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
Chest pain > 2 weeks	No	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	Yes	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p

Flu-like symptoms in past 2 weeks	No	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	Yes	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
	Not recorded	n (%)	x (%)	x (%)			x (%)			x (%)		
WHO symptom screen	Negative	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	Positive	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
Number of TB symptoms	0	n (%)	x (%)	x (%)	1		x (%)	1		x (%)	1	
	1	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
	2	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
	3	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p
	4	n (%)	x (%)	x (%)	x (x–x)	p	x (%)	x (x–x)	p	x (%)	x (x–x)	p

OR = Odds ratio

4.6 PRIMARY AND SECONDARY OUTCOMES TABLES (BINARY ANALYSIS)

4.6.1 CUMULATIVE TEST PERFORMANCE FOR PRIMARY END-POINT (TWO-SAMPLE) PREVALENT AND INCIDENT TB CASES DURING FOLLOW-UP (PRIMARY AIM)

4.6.2 CUMULATIVE TEST PERFORMANCE FOR COMBINED PRIMARY AND SECONDARY END-POINT (ONE- OR TWO-SAMPLE) PREVALENT AND INCIDENT TB CASES DURING FOLLOW-UP (PRIMARY AIM)

4.6.3 SCREENING/TRIAGE TEST PERFORMANCE FOR PRIMARY END-POINT (TWO-SAMPLE) PREVALENT TB CASES AT ENROLMENT (SECONDARY AIM 1)

4.6.4 SCREENING/TRIAGE TEST PERFORMANCE FOR COMBINED PRIMARY AND SECONDARY END-POINT (ONE- OR TWO-SAMPLE) PREVALENT TB CASES AT ENROLMENT (SECONDARY AIM 1)

4.6.5 PREDICTIVE TEST PERFORMANCE FOR PRIMARY END-POINT (TWO-SAMPLE) INCIDENT TB CASES DURING FOLLOW-UP

4.6.6 PREDICTIVE TEST PERFORMANCE FOR COMBINED PRIMARY AND SECONDARY END-POINT (ONE- OR TWO-SAMPLE) INCIDENT TB CASES DURING FOLLOW-UP

	Missing or indeterminate result	TP	FN	TN	FP	AUC, % (95% CI)	RR or CIR, n (95% CI)	Sensitivity‡ % (95% CI)	Specificity‡ % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)	NNS, n (95% CI)
Symptom-based screening*												
1 or more WHO symptoms positive	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
2 or more WHO symptoms positive	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
3 or more WHO symptoms positive	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
4 or more WHO symptoms positive	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
Cough for ≥2 weeks (with or without presence of other symptoms)	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
Cough for ≥2 weeks and at least one other symptom	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
Any symptom (including chest pain and flu-like symptoms)	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
QuantiFERON-TB Gold In-tube assay (QFT)												
Optimal QFT threshold (≥xx IU/ml) ‡	x	x	x	x	x	x% (x–x)	x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
A-priori QFT threshold, as per manufacturer (≥0.35 IU/ml)	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
QFT threshold >0.7 IU/ml+	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
QFT <0.2 IU/ml NEGATIVE+ QFT 0.2–0.7 IU/ml INDETERMINATE+ QFT >0.7 IU/ml POSITIVE+	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
QFT threshold ≥1 IU/ml§	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
QFT threshold ≥4 IU/ml§	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
Urine Alere LAM	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
COR (RISK11) †												
Optimal COR threshold (≥xx) ‡	x	x	x	x	x	x% (x–x)	x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
COR threshold ≥90	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
COR threshold ≥80	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
COR threshold ≥70	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
COR threshold ≥60	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
COR threshold ≥50	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)

COR threshold ≥40	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
COR threshold ≥30	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
COR threshold ≥20	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
COR threshold ≥10	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
COR (RISK11): indeterminate results included												
Optimal COR threshold (≥xx) ‡	x	x	x	x	x	x% (x–x)	x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
COR threshold ≥90	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
COR threshold ≥80	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
COR threshold ≥70	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
COR threshold ≥60	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
COR threshold ≥50	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
COR threshold ≥40	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
COR threshold ≥30	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
COR threshold ≥20	x	x	x	x	x		x (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
COR threshold ≥10	x	x	x	x	x		x (x–x)	x% (x–x)	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; CIR = Cumulative incidence ratio; RR = Relative risk; 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.

* TB symptoms include loss of weight, persistent unexplained cough, fever, or night sweats for longer than two weeks; or any haemoptysis.

[+] Nemes E, Rozot V, Geldenhuys H, Bilek N, Mabwe S, Abrahams D, et al. Optimization and Interpretation of Serial QuantiFERON Testing to Measure Acquisition of Mycobacterium tuberculosis Infection. American Journal of Respiratory and Critical Care Medicine. 2017;196(5):638-48. doi: 10.1164/rccm.201704-0817OC.

[§] Winje BA, White R, Syre H, Skutlaberg DH, Oftung F, Mengshoel AT, et al. Stratification by interferon-gamma release assay level predicts risk of incident TB. Thorax. 2018. doi: 10.1136/thoraxjnl-2017-211147.

[†] Indeterminate results excluded.

[‡] Threshold with specificity benchmarked at 70% for screening/triage and 75% for prediction of incident TB cases.

4.7 PERFORMANCE OF THE COR (RISK11) AND QFT FOR IDENTIFICATION OF TB DISEASE OVER A 15-MONTH PERIOD, STRATIFIED BY THE TIME INTERVAL TO DISEASE (TIME-DEPENDENT ANALYSIS)

4.7.1 CUMULATIVE TEST PERFORMANCE FOR PRIMARY END-POINT (TWO-SAMPLE) PREVALENT AND INCIDENT TB CASES DURING FOLLOW-UP (PRIMARY AIM)

4.7.2 CUMULATIVE TEST PERFORMANCE FOR COMBINED PRIMARY AND SECONDARY END-POINT (ONE- OR TWO-SAMPLE) PREVALENT AND INCIDENT TB CASES DURING FOLLOW-UP (PRIMARY AIM)

4.7.3 PREDICTIVE TEST PERFORMANCE FOR PRIMARY END-POINT (TWO-SAMPLE) INCIDENT TB CASES DURING FOLLOW-UP (SECONDARY AIMS 2 AND 3)

4.7.4 PREDICTIVE TEST PERFORMANCE FOR COMBINED PRIMARY AND SECONDARY END-POINT (ONE- OR TWO-SAMPLE) INCIDENT TB CASES DURING FOLLOW-UP (SECONDARY AIMS 2 AND 3)

Test and time interval to TB disease	Cuminc COR+, % (95% CI)	Cuminc COR-, % (95% CI)	CIR, n (95% CI)	p value	AUC, % (95% CI)	Sensitivity‡ % (95% CI)	Specificity‡ % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)	NNS, n (95% CI)
QuantiFERON-TB Gold In-tube assay (QFT) †										
0 to 3 months	x (x–x)	x (x–x)	x (x–x)	x	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
0 to 6 months	x (x–x)	x (x–x)	x (x–x)	x	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
0 to 12 months	x (x–x)	x (x–x)	x (x–x)	x	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
0 to 15 months	x (x–x)	x (x–x)	x (x–x)	x	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
COR (RISK11) †										
0 to 3 months	x (x–x)	x (x–x)	x (x–x)	x	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
0 to 6 months	x (x–x)	x (x–x)	x (x–x)	x	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
0 to 12 months	x (x–x)	x (x–x)	x (x–x)	x	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
0 to 15 months	x (x–x)	x (x–x)	x (x–x)	x	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)

Cuminc = Cumulative incidence; CIR = Cumulative incidence ratio; 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.

† Indeterminate results excluded.

‡ Threshold with specificity benchmarked at 75% for prediction of incident TB cases.

5 FIGURES

5.1 TYPES OF FIGURES

5.1.1 Test performance: ROC curves (sensitivity versus 100-specificity)

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

5.1.3 Test score distribution: violin/box-and-scatter plots (Prevalent and incident TB cases versus TB negative controls)

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

5.2 ENDPOINTS FOR FIGURES

5.2.1 Cumulative test performance for primary endpoint (two-sample) prevalent and incident TB cases during follow-up (Primary aim)

5.2.2 Cumulative test performance for combined primary and secondary endpoint (one- or two-sample) prevalent and incident TB cases during follow-up (Primary aim)

5.2.3 Screening/triage test performance for primary endpoint (two-sample) prevalent TB cases at enrolment (Secondary aim 1)

5.2.4 Screening/triage test performance for combined primary and secondary endpoint (one- or two-sample) prevalent TB cases at enrolment (Secondary aim 1)

5.2.5 Predictive test performance for primary endpoint (two-sample) incident TB cases during follow-up (Secondary aim 2)

5.2.6 Predictive test performance for combined primary and secondary endpoint (one- or two-sample) incident TB cases during follow-up (Secondary aim 2)

5.3 TESTS

5.3.1 COR (RISK11)

5.3.2 QuantiFERON

6 APPENDIX

6.1 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%

6.2 CONCOMITANT MEDICATION CODING

MEDICATION NAME	ART	IPT	VIT B6	TB RX	OTHER
(INHALER) ASTHAVENT					Y
2.5 SELSUN					Y
3TC+AZT	Y				
ABACAVIR	Y				
ABACAVIR / LAMIVUDINE	Y				
ABACAVIR LAMIVUDINE	Y				
ABC/3TC	Y				
ACC 200					Y
ACRIPTAZ	Y				
ADCO EFAVIRENZ	Y				
ADCO-EMTERVIR	Y				
ADCO-LAMIVUDINE/ ZIDOVUDINE	Y				
ADCODOL					Y
ADRENALINE					Y
AEVS FIXED DOSE COMBINATION	Y				
ALCOPHYLLEX					Y
ALLERGEX					Y
ALLUVIA	Y				
ALTROIZA	Y				
ALTROIZA (TDF/FTC\EFC)	Y				
ALUVIA	Y				
ALUVIN	Y				
AMITRIPTILLIEN					Y
AMITRIPTILLINE					Y
AMITRIPTYLINE					Y
AMLODIPINE					Y
AMOXICILLIN					Y
AMOXICYCLIN					Y
AMOXICYLLIN					Y
AMOXIL					Y
AMOXYCILIN					Y
AMOXYCILLIN					Y
AMOXYCILLIN CLAVULINIC ACID					Y
AMOXYCLLIN					Y
AMOXYLLIN					Y
AMPCILLIN					Y
AMTAS					Y
ANOXICILLIN					Y
ANTI HYPER TENSIVE ENALAPRIL					Y
ANTIBIOTIC (UNKOWN)					Y
ANTRIPLA	Y				
ANUSOL					Y
AQEOUS CREAM					Y
AQUEOUS CREAM					Y
ART (FDC)	Y				
ART FDC	Y				
ARTROIZA	Y				
ARV	Y				
ARV (FDC)	Y				
ARV'S (FDC)	Y				
ARV'S FDC	Y				
ARV(FDC)	Y				
ARVS (FDC)	Y				
ASHTAVENT					Y
ASTHAVENT					Y
ATANEF	Y				
ATAZA	Y				
ATAZANAVIR	Y				
ATAZOR	Y				
ATENEF	Y				
ATENEF (FDC)	Y				
ATENET	Y				
ATIZAVIRE	Y				
ATOIZA	Y				
ATRAZA	Y				
ATRIPLA	Y				
ATRIZA	Y				
ATROIZA	Y				
ATROIZA (FDC)	Y				
ATROIZA FDC	Y				
ATROIZIA	Y				
ATROZIA	Y				
AUGMENTIN					Y
AUGMNTIN					Y

AUGUMENTIN					Y
AUSTELL PARACETAMOL					Y
AZITHRO					Y
AZITHROMYCIN					Y
AZITHROMYSN					Y
AZT/3TC	Y				
AZT\3TC	Y				
BABALAS REMEDY					Y
BACTRIM					Y
BECLATE					Y
BETADINE					Y
BETAGESIC					Y
BRUFEN					Y
BUDEFLAM					Y
BUDESONIDE					Y
BURFEN					Y
CA GLUCOMAL					Y
CALCIUM GLUCENT					Y
CALGITROL					Y
CALIUM CHLORIED IN 200ML NACL					Y
CALIUM CHLORIED IN 200ML NACL					Y
CAMBIVIR	Y				
CARBAMAZEPINE					Y
CARBAMAZEPINE (TEGRETOL)					Y
CARBIMAZOLE					Y
CEFTRIAXONE					Y
CHLOPHENIRAMINE MALEATE					Y
CHLOROPHERAMINE					Y
CHLORPHENARIMINE MALEATE					Y
CHLORPHENIRAMINE					Y
CHLORPHENIRAMINE MALEATE					Y
CHLORPHERAMINE					Y
CHRORPHENIRAMINE MALEATE					Y
CIFRAN					Y
CIPROBAY					Y
CIPROFLOX					Y
CITALOPRAM					Y
CLAXACILLIN					Y
CLEXANE					Y
CLOFAZIMINE				Y	
CLOMTRIMAZOLE					Y
CLOTRIMAZOLE					Y

CO AMOXY CLAN					Y
CO TRIMOXAZOLE					Y
CO-TRIMOXAZOLE					Y
CORTIMOXAZOLE					Y
COTRANAZOLE					Y
COTRIMOXAZOLE					Y
COTROMOXAZOLE					Y
COZOLE					Y
CR500 EPILIM					Y
DEPO PROVERA					Y
DEPO-PROVERA					Y
DEPO. PROVERA					Y
DEPOT PROVERA					Y
DEPRO PROVERA					Y
DIAZEPAM					Y
DIPHENHYDRAMINE					Y
DIPHENHYDRAMINE HYDROCHLORIDE					Y
DISPRIN					Y
DOXYCYCLINE					Y
DUMIVA	Y				
DUROBAC					Y
DYNASPOR					Y
EFARAVINS	Y				
EFAVARENCE	Y				
EFAVARENZ	Y				
EFAVARENZE	Y				
EFAVERENZ	Y				
EFAVIRENCE	Y				
EFAVIRENE	Y				
EFAVIRENZ	Y				
EFAVIRENZ/ EMTRICITABINE/ TENOFOVIR	Y				
EFAVIRENZE	Y				
EFAVIREZ	Y				
EFEVARENZ	Y				
EFIVARENS	Y				
EFIVARENZ	Y				
EFV	Y				
ELTROXIN					Y
EMCTRITABINE	Y				
EMICITRICITABINE	Y				
EMITIRICATIBANE	Y				

EMITRACITABINE	Y				
EMITRICIBALE	Y				
EMITRICIBANE	Y				
EMITRICIBATE	Y				
EMITRICIBIN/TDF/EFV	Y				
EMITRICITABANE	Y				
EMITRICITABATE	Y				
EMITRICITABIN	Y				
EMITRICITABINE	Y				
EMITRICITIBINE	Y				
EMTRICIBATINE	Y				
EMTRICITABIEN	Y				
EMTRICITABINE	Y				
EMTRICITATABINE	Y				
ENALAPRIL					Y
ENUTRICUTABINE	Y				
ENVAS					Y
ENVASLO					Y
EPILIM					Y
EPILIUM CR					Y
ERTAPENEM					Y
ERYTHROMYCIN					Y
ESONIAZID		Y			
ETC	Y				
ETENOGEREL					Y
ETHAMBUTOL				Y	
ETHIONAMIDE				Y	
ETONOGESTREL					Y
FDC	Y				
FDC (TRNOFIVIR, EFAVERENZ, EMTRICITABINE)	Y				
FDC (ARVS)	Y				
FDC (TDF/EMC/EFV)	Y				
FDC ART	Y				
FDC ARV	Y				
FDC ARV (TRIBUSS)	Y				
FDC ARV TRIBUSS	Y				
FDC ARVS	Y				
FDC ATROIZA	Y				
FDC TESOFIEL	Y				
FDC TRIBUSS	Y				
FE 504					Y
FERROUS SULPHATE					Y

FIXED DOSE (TDF, FTC,EFV)	Y				
FIXED DOSE ANTI RETROVIRAL THERAPY	Y				
FIXED DOSE ARV'S	Y				
FIXED DOSE COMBINATION ARVS	Y				
FIXED DOSE COMBINATION (TDF, EMB, EFV)	Y				
FIXED DOSE COMBINATION ART	Y				
FIXED DOSE COMBINATION ARV	Y				
FLAGYL					Y
FLUCLOXICALLIN					Y
FLUCONAZOLE					Y
FLUOXETINE					Y
FLUPAYNE					Y
FLUTEX					Y
FOLIC ACID					Y
FTC	Y				
GENTAMICIN					Y
GRAND-PA					Y
HCTZ					Y
HYDROCHLOROTHIAZIDE					Y
HYDROCHLOROTHYAZIDE					Y
HYDROCHROTHIAZIDE					Y
HYDROCURNICHE					Y
HYPACE					Y
IBUNATE					Y
IBUPROFEN					Y
ILVITRIM					Y
IMPLANON					Y
IMPLANT					Y
INDO AMOXYCILLIN					Y
INH		Y			
INSONIAZIDE		Y			
IONIAZID		Y			
IRON SUPPLEMENTATION					Y
ISINIAZID		Y			
ISIONAZID		Y			
ISIONIAZID		Y			
ISONAIZED		Y			
ISONAIZID		Y			
ISONAZIAD		Y			
ISONIAID		Y			
ISONIAZED		Y			

ISONIAZIAD		Y			
ISONIAZID		Y			
ISONIAZIDE		Y			
ISONIZAZID		Y			
IWH		Y			
IZONAZAID		Y			
IZONIAZIDE		Y			
KANAMYCIN				Y	
KAVIMUN	Y				
KIVEXA	Y				
LACSON					Y
LACTULOSE					Y
LAMIVADINE, ZIDOVIDINE	Y				
LAMIVIDUNE	Y				
LAMIVUDENE / IDOVUDINE	Y				
LAMIVUDINE	Y				
LAMIVUDINE / ZIDOVUDINE	Y				
LAMIVUDINE AND ZIDOVUDINE	Y				
LAMSOPRAZOLE					Y
LAMUVIDINE	Y				
LAMZID EFAVIRENZ	Y				
LANSELOC					Y
LANSOLOC					Y
LAPINAVIR	Y				
LASIX					Y
LEROFLXICIN				Y	
LIPONAVIR	Y				
LIQUID PARAFIN					Y
LOPERAMIDE					Y
LOPINAVIR	Y				
LOPINAVIR / RITONAVIR	Y				
LOPINAVIR/ RITONAVIR	Y				
LOPINAVIR/ RITORAVIR	Y				
LOPINOVIR	Y				
LORAZEPAM					Y
LOSEC					Y
LUMIVUDINE	Y				
LUSIX					Y
MAXALON					Y
MED LEMON					Y
MEDROXY PROGESTERONE ACETATE					Y
MEDROXYPROGE STERONE ACETATE					Y

MEDROXYPROGESTERONE					Y
MEDROXYPROGESTERONE ACETATE					Y
MEDROXYPROGESTRONE ACETATE					Y
METOCLOPRAMIDE					Y
METOCLOPROMIDE					Y
METROCLOPRAMIDE					Y
METRONIDAZOLE					Y
MGS04 IN 200ML NACL					Y
MIST POT CHLOR					Y
MORPHINE					Y
MOXIFLAXACIN					Y
MOXIFLOXACIN				Y	
MOXYMAX					Y
MULTIVIT					Y
MULTIVITAMIN					Y
MULTIVITS (VIT B6)					Y
MVT					Y
NAPAMOL					Y
NEVARAPINE	Y				
NEVIRAPIN	Y				
NEVIRAPINE	Y				
NORDETTE					Y
NORETHISTERONE ENANTHATE					Y
NORETHISTHERONE ENANTHATE					Y
NORGESTREL/ ETHINYL OESTRADIOL					Y
NORMAL SALINE					Y
NOZER OMEPRAZOLE					Y
NU ISTERATE					Y
NU-ISTERATE					Y
NUCOTRIM					Y
NUR ISTERATE					Y
NUR ISTRATE					Y
NUR-ISTERATE					Y
NUR-ISTRATE					Y
NURISTERATE					Y
NURISTRATE					Y
NYSTATIN					Y
ODIMUNE	Y				
OMEPRAZOLE					Y
OMEPROZOLE					Y
ORALCON					Y
OVARAL					Y

OVRAL					Y
OXYMETALOLINE					Y
OXYMETAZOLINE HYDRACHLORIDE					Y
PAINBLOCK					Y
PAINBLOK					Y
PANADO					Y
PARACETAMOL					Y
PARACETEMOL					Y
PERTOGEN					Y
PETEGEN					Y
PETOGEN					Y
PHARMAPRESS					Y
PHENERGAN					Y
PHOLCODINE					Y
PIRIDOXINE			Y		
POTASSIUM CHLORIED					Y
PREDISONE					Y
PREDNISONE					Y
PROPAN					Y
PURBAC TABS					Y
PYIRIDOXINE			Y		
PYNDOXINE			Y		
PYRAZANIMIDE				Y	
PYRAZINAMIDE				Y	
PYREDOXIN			Y		
PYRID0XINE			Y		
PYRIDIXINE			Y		
PYRIDOXIDE			Y		
PYRIDOXIN			Y		
PYRIDOXINE			Y		
PYRIDOXINE (VITAMIN B)			Y		
PYRIDOXIXE			Y		
PYRIODOXINE			Y		
PYRIXODINE			Y		
PYRODOXCIN			Y		
PYRODOXINE			Y		
PYRYDOXINE			Y		
RANITIDINE					Y
REFAFOUR E-275				Y	
REPIVATE					Y
RESPIRADONE					Y
RETONOVIR	Y				

RIDA2					Y
RIDAQ					Y
RIFAFOUR				Y	
RIFAFOUR E-275				Y	
RIFINAH				Y	
RIMACTANE		Y			
RINGERS LACTATE					Y
RISPERIDONE					Y
RITENOVIR	Y				
RITONAVIR	Y				
RITONOVIR	Y				
RIZENE	Y				
ROCEPHIN					Y
SALBUTAMOL (ASTHAVENT)					Y
SALTERPYN					Y
SELSYN 2-5/.					Y
SIMVASTATIN					Y
SIRTURO					Y
SONKE EFAVIRENZ	Y				
STAVUDINE	Y				
STILPAYNE					Y
SULFAMETHOXAZOLE					Y
SULPHA TRIMETH					Y
TDF	Y				
TDF \ EFV\ ETC	Y				
TDF, FTC, EFV	Y				
TDF,EMT,EFV	Y				
TDF/ EFV/ FTC	Y				
TDF/ EMT/EFV	Y				
TDF/ FCT/EFV	Y				
TDF/DTC/EFV	Y				
TDF/EFV/ETC	Y				
TDF/EFV/FTC	Y				
TDF/EMT/EFV	Y				
TDF/EMT/EFV (ART)	Y				
TDF/ETB/EFV	Y				
TDF/ETD/EFV	Y				
TDF/FTC	Y				
TDF/FTC/EFC	Y				
TDF/FTC/EFV	Y				
TDF\EMI\EFV	Y				
TDF\EMT\EFV	Y				

TDF\ETB\EFV	Y				
TDF\ETC\EFV	Y				
TDF\EVNT\EFV	Y				
TDF\FTC\EFV	Y				
TENEMINE	Y				
TENO/EMTR/EFA	Y				
TENOFEVIR	Y				
TENOFIVER	Y				
TENOFIVIR	Y				
TENOFORVIR	Y				
TENOFOUIR	Y				
TENOFOVIR	Y				
TENOFOVIR/ EMTRICITABINE/EFAVIRENZ	Y				
TENOVIR	Y				
TENOVOFIR	Y				
TERIZIDONE				Y	
THEOPHLLINE					Y
THIAMINE					Y
THYROXINE					Y
TRAMADOL					Y
TRAMADUL					Y
TRAPHASIL					Y
TRIBUS	Y				
TRIBUS (TDF/EFT/EFV)	Y				
TRIBUS (TDF, FTC. EFV)	Y				
TRIBUS (TDF/EMT/EFV)	Y				
TRIBUS (TDF/FTC/EFN)	Y				
TRIBUS (TDF/FTC/EFV)	Y				
TRIBUS (TDF\FTC\EFN)	Y				
TRIBUS (TDF\FTC\EFV)	Y				
TRIBUS. (TDF/FTC/EFV	Y				
TRIBUS(TDF\FTC\EFV)	Y				
TRIBUS(TDF|EMT|EFV)	Y				
TRIBUS(TDF|FTC|EFV)	Y				
TRIBUSS	Y				
TRIBUSS (TDF,3TC,ETC)	Y				
TRIBUSS (TDF/EDC/EFV)	Y				

TRIBUSS (TDF/EMT/EFV)	Y				
TRIBUSS (TDF/FTC/EFV)	Y				
TRIBUSS (TDF\3TC/EFV)	Y				
TRIPHASIL					Y
TRISUS (TDF\FTC\EFV)	Y				
TRIVENZ	Y				
TRIXAZOLE					Y
TRIZANOLE					Y
TRUVADA	Y				
UNKNOWN ART	Y				
UNKNOWN ARV	Y				
UNKNOWN PAIN KILLERS					Y
UNKNOWN SINGLE DOSE ART REGIMENT	Y				
UNKNOWN SINGLE DOSE. ARV TABLETS	Y				
VASTOR					Y
VENTEZE					Y
VIIT C					Y
VIT B					Y
VIT B6					Y
VIT BCO					Y
VIT C					Y
VITAMIN B COMPLEX					Y
VITAMIN B-CO					Y
VITAMIN B6					Y
VITAMIN BCO					Y
VITAMIN C					Y
VITB					Y
VITB 6			Y		
VOLTAREN					Y
WINTROP ISONIAZIDE		Y			
ZETENOL					Y
ZIDOMAT	Y				
ZIDOVUDINE	Y				
ZIDOVUDINE/LAMIVUDINE	Y				
ZIOLOVUDINE	Y				
ZIVOLAM	Y				
ZOVILAM	Y				

7 REFERENCES

Aalen O. 1978. Nonparametric inference for a family of counting processes. *The Annals of Statistics.*701-726.

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.

Carpenter J, and Bithell J. 2000. Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. *Statistics in Medicine.* 19 (9):1141-1164. doi: 10.1002/(sici)1097-0258(20000515)19:9<1141::Aid-sim479>3.0.Co;2-f.

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.

Darboe F, Mbandi SK, Naidoo K, Yende-Zuma N, Lewis L, Thompson EG, Duffy FJ, Fisher M, Filander E, van Rooyen M*, et al.* 2019. Detection of Tuberculosis Recurrence, Diagnosis and Treatment Response by a Blood Transcriptomic Risk Signature in HIV-Infected Persons on Antiretroviral Therapy. *Front Microbiol.* 10:1441. doi: 10.3389/fmicb.2019.01441.

DeLong ER, DeLong DM, and Clarke-Pearson DL. 1988. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics.* 44 (3). doi: 10.2307/2531595.

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.

Koopman P. 1984. Confidence intervals for the ratio of two binomial proportions. *Biometrics.*513-517.

Nam JM. 1995. Confidence limits for the ratio of two binomial proportions based on likelihood scores: non-iterative method. *Biometrical Journal.* 37 (3):375-379.

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.

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

Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, and Muller M. 2011. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics.* 12:77. doi: 10.1186/1471-2105-12-77.

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.