PredictTB Trial

NIAID # 16-I-N133

Using Biomarkers to Predict TB Treatment Duration

Statistical Analysis Plan

Trial Statistician	Principal Investigator
Lori E Dodd, PhD Biostatistics Research Branch Division of Clinical Research, NIAID 5601 Fishers Lane, Room 4C31 Mail Stop Code: 9820 Bethesda, MD 20892 Phone: 240-669-5247 Email: doddl@mail.nih.gov	Clifton E. Barry III, Ph.D. National Institute for Allergy and Infectious Diseases (NIAID), Laboratory of Clinical Infectious Diseases (LCID), Tuberculosis Research Section (TRS), Building 33, Room 2w10 Bethesda, MD 20892 Phone: (301) 693-4665 Email: cbarry@niaid.nih.gov

Enrollment sites:

Site Name	**Country**	**City**
Henan Provincial Chest Hospital (HPCH)	China	Zhengzhou, Henan Province
Kaifeng City Institute of Tuberculosis Prevention and Control	China	Kaifeng, Henan Province
Xinmi City Center for Disease Control and Prevention	China	Xinmi, Henan Province
Xinxiang City Institute of Tuberculosis Prevention and Control	China	Xinxiang, Henan Province
Zhongmu County Station for Disease Control and Prevention	China	Zhongmu, Henan Province
University of Cape Town (UCT) Lung Institute	South Africa	Cape Town
Stellenbosch University	South Africa	Tygerberg
TASK Applied Science, Inc	South Africa	Bellville/Delft
Khayelitsha Site B	South Africa	Khayelitsha
South African Tuberculosis Vaccine Initiative (SATVI)	South Africa	Cape Town/Worcester

1 Introduction

PredictTB is a randomized controlled trial evaluating the PredictTB treatment-shortening criteria (a combination of imaging, GeneXpert and treatment adherence results). The study will prospectively identify patients at low risk based on their baseline radiographic extent of disease, and further refine this risk score by evaluating the rate of resolution of the lung pathology (CT) and inflammation (PET) at one month as well as checking an end-of-treatment GeneXpert MTB/RIF test for the sustained presence of bacteria. Patients classified as low risk will be randomized to receive a shortened 4-month or a full 6-month course of therapy. If successful, this trial will both offer a badly needed alternative to culture status as a trial-level surrogate marker for outcome as well as provide critical information for preclinical and early clinical efforts to identify new agents and combinations with the potential to shorten therapy.

2 Design and objectives

This is a prospective, randomized, phase 2b noninferiority trial in pulmonary drug suscpetible-TB participants. Eligible and consented participants will start on standard treatment ("HRZE") and undergo evaluations to determine eligibility for treatment shortening according to the PredictTB criteria (Table 1). Participants who meet the PredictTB criteria will be randomized at week 16 to arm B (standard-duration treatment) or arm C (shortened treatment). All participants will be followed for 72 weeks (18 months) following treatment initiation.

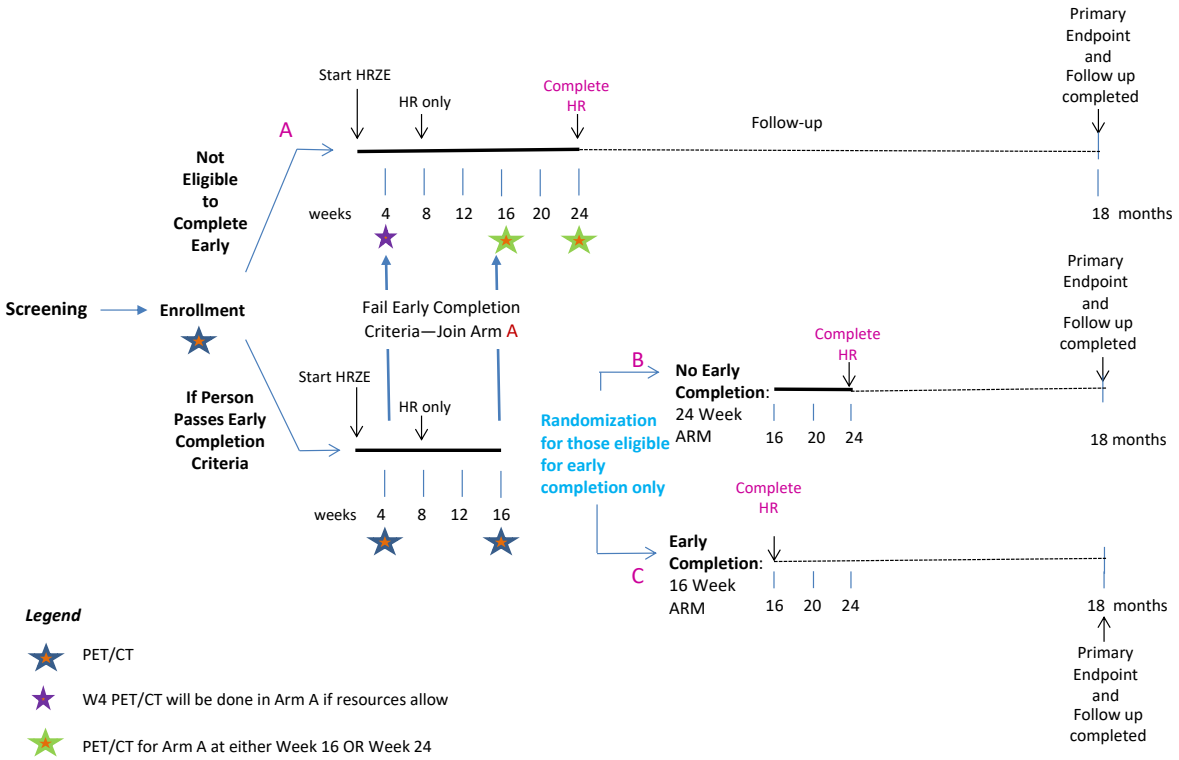

Figure 1: Study Flow

Table 1: Early Treatment Completion Criteria: All early treatment criteria must be met for the participant to be eligible for randomization

Early Completion criteria:	**Determined at Week 16 – unless known to have failed a radiographic criterion at baseline or week 4.**
Radiographic criteria	Baseline PET/CT: • No total lung collapse of a single side, AND • No single cavity air volume on CT scan >30 mL, AND • CT scan hard volume (-100 to +100 HU density) <200 mL, OR • PET total activity <1500 units Week 4 PET/CT: • All individual cavities decrease by >20% (unless cavity <2 mL), AND • CT scan hard volume does not increase by >10% unless the increase is <5 mL, OR • PET total activity does not increase by >30% unless the increase is <50 units
Bacterial load criterion	Week 16 Xpert MTB/RIF cycle threshold ≥28*
Adherence criterion	Minimum of 100 out of 112 doses received by week 16 (Sec 5.2)

2.1 Study objectives and hypotheses

Primary Objective

To demonstrate that the 72 week poor outcome proportion under shortened treatment at week 16 (Arm C) is not inferior (NI margin=0.07) to that from standard-duration treatment (stopped at week 24; Arm B), in participants classified as low risk for disease failure and relapse the PredictTB criteria.

Primary hypothesis: Amongst participants who satisfy the PredictTB criteria, the proportion of poor outcomes will be similar between participants undergoing 16 and 24 weeks of standard treatment.

Secondary Objectives

1) The compare the proportion of poor outcomes from low-risk participants with shortened treatment to that of a representative 6-month standard-of-care population.
   Hypothesis: The poor outcome rate among low-risk participants with shortened treatment will be like that of a representative 6-month standard of care population.

2) To evaluate the association of demographic, radiographic, bacterial load, microbiologic, and immunologic markers (at baseline and during treatment) for predicting treatment failure in the following participant cohorts:
    - Pooling arms A and B, participants receiving the same duration of therapy, to evaluate the risk criteria.
    - Between arms B and C, to evaluate whether there are any covariates which predict greater rates of failure under treatment shortening.

    Hypothesis: In univariate and multivariate analyses, demographic, radiographic, bacterial load, microbiologic, and immunologic markers are associated with treatment failure.

3) To evaluate the association of demographic, radiographic, bacterial load, microbiologic, and immunologic markers (at baseline and during treatment) for predicting subsequent relapse in the following participant cohorts:
    - Pooled arms A and B (i.e., participants receiving the same duration of therapy) to evaluate the risk criteria.
    - Between arms B and C, to evaluate whether there are any covariates which predict greater rates of relapse under treatment shortening.

    Hypothesis: In univariate and multivariate analyses, demographic, radiographic, bacterial load, microbiologic, and immunologic markers are associated with subsequent relapse.

4) To compare the ability of bacterial load (TTP and Xpert MTB/RIF cycle threshold) at different time points to distinguish between poor and good outcomes.

    Hypothesis: Bacterial load markers (TTP and Xpert MTB/RIF cycle threshold) collected at later time points are better markers of poor outcomes than markers collected at earlier times.

5) To evaluate sub-breakpoint MICs as a significant predictor of poor outcomes in 16 and 24 week treatment regimens in a subset of participants.

    Hypothesis: Pharmacokinetic and sub-breakpoint MIC measurements are associated with poor outcomes in 16 and 24 week treatment regimens.

2.2 Study visit schedule

	Screening	D0	W1[V] (D7)	W2 (D14)	W4 (D28)	W8[D] (D56)	W12 (D84)	At Week 16	W16 (D112)	W20 (D140)	W24 (D168)	W36 (D252)	W48 (D336)	W60 (D420)	W72 (D504)	TB Recurrence
Main Study Informed Consent (Plus Genetic and HIV in RSA)	X							RANDOMIZE to Arm B or C								
Medical History/Focused History	X	X	X	X	X	X	X		X	X	X	X	X	phone call	X	X
Physical Exam/Focused Physical Exam	X	X	X	X	X	X	X		X	X	X	X	X		X	X
Sputum Collection[ξ]																
Smear/Culture	XX[H]	(X)[S]	X	X	X	X	X	then	X	X	X	X	X		X	X
GeneXpert	X	(X)[S]				X			X[E]		X				X	X
Biomarkers		X		X	X[g]	X[g]			X[g]		X[g]		X		X	X[g]
Saliva		X			X				X		X					X
Blood collection																
CBC/Chems/LFT[G]	X							and if not assigned to Arm A,								X
Biomarkers		X	X	X	X	X			X		X		X		X	X
Pregnancy Test[P] (serum at screening, urine for others)	x[h](serum)															
HIV Testing	X															X
Plasma drug levels										X						
Finger Stick		X			X[%]				Arms B/C and randomized A		Randomized for scan in Arm A only					X
Urine collection																
Biomarkers		X	X	X	X	X		Review treatment completion criter	X		X		X		X	X
Pregnancy Test[P] (serum at screening, urine for others)		X			X[%]				Arms B and C only		Arms A and B only					X
CXR	X								Arm C only		Arms A and B only					X
FDG-PET/CT[W]		X			X[%]				Arms B/C and randomized A		Randomized for scan in Arm A only					X
Adherence monitoring			X	X	X	X	X		X	Arms A and B only	Arms A and B only					

[ξ] Additional sputum may be collected if contaminated or otherwise compromised A subject could be called back for this additional sputa collection, if necessary.

[H] Sputum at screening will also be used for screening Xpert

[S] Day 0 culture/smear/Xpert only done if NOT within 7 days of screening, or per site's procedures.

[h] Pregnancy testing from screening may be used for the D0 PET/CT scan if D0 is within 2 days of the screen

[G] These will be performed at any visit if clinically significant.

[P] Before any PET/CT scan or CXR is done, a pregnancy test will be done for applicable females. If the pregnancy test is positive, the PET/CT scan will not be performed. The CXR may still be performed with appropriate shielding.

[E] for those eligible for randomization

[D] At week 8, ethambutol and pyrazinamide will be discontinued

*[W] **PET/CT scan windows** : baseline w/i 7 days after treatment initiation; W4 must be at least 4 wks after baseline scan with a -3/+7 d window; W16 and 24 scan w/i 14 d of visit; relapse ASAP, but w/i 2 wks of recurrence*

*[V] **Visit windows:** Week 1-2: +/- 3 days; Week 4-24: +/- 7 d, noting that Weeks 16 and 24 should be as close as possible to actual date; Week 36-72: +/- 30 days.*

[g] If sputum is not available for biomarkers, it will not be a protocol deviation.

[%] Arms B and C will have PET/CT (with prior finger stick and pregnancy test). Arm A will have if logistically possible.

2.3 Randomization.

Participants who satisfy the PredictTB criteria at week 16 will be randomized to stop treatment immediately or continue for the standard duration. Randomization is stratified by site, using random block sizes of four and six, resulting in nine separate randomization files. Participant randomization will occur through the DataFax system, as documented in the Data Management Plan.

Participants who are considered high-risk for relapse (i.e., Arm A) will be randomized to obtain a third image at either week 16 or week 24. This randomization is stratified according to when the subjects move to arm A (i.e., baseline, week 4 or week 16). Participant randomization will occur through the DataFax system.

2.4 Sample Size Considerations

For this study, the sample size is calculated for Arms B and C, which are used for the primary endpoint. Because these are lower risk participants, we expect a treatment success rate of 97%, corresponding to a poor outcome rate of 3%. Table 2 provides power calculations for demonstrating noninferiority (with NI margin 0.07) using a 90% confidence interval with sample sizes of 129 and 155 per group, adjusted for a 10% loss to follow-up rate. With true success rates of 97% in both arms, study power is greater than 90% with only 129 participants per group. However, to increase power to accommodate a scenario in which the true success rate in the four-month treatment arm is slightly lower than the six-month arm, a sample size of 155 per treatment arm was selected. We expect that approximately 50% of participants will be classified as higher risk and be placed into Arm A, giving a total study sample size of 620 participants. In the event that the proportion of participants considered lower risk by our treatment completion criteria is less than 50%, participant enrollment will continue until 155 participants per arm are enrolled into Arms B and C and at least 200 participants are enrolled into Arm A. Note that the power calculations were based on an exact binomial CI of the difference in poor outcomes, as an approximation to the primary analysis, which will be based on the Kaplan-Meier based estimates of poor outcomes at week 72.

Table 2. Power calculations for total sample sizes of 129 and 155 per group (arms B and C) for different poor outcome rates across and between treatment arms.

Poor outcome rate by study arm		Power for concluding NI with 7% margin, 5% type I error rate, and 10% loss to follow-up	
6-month tx	4-month tx	Sample size 129 per group	Sample size 155 per group
0.01	0.01	0.999	1

0.01	0.02	0.984	0.994
0.01	0.03	0.863	0.912
0.02	0.02	0.985	0.994
0.02	0.03	0.903	0.942
0.02	0.04	0.726	0.792
0.03	0.03	0.932	0.963
0.03	0.04	0.803	0.862
0.03	0.05	0.621	0.689
0.04	0.04	0.862	0.911
0.04	0.05	0.716	0.782
0.04	0.06	0.545	0.609
0.05	0.05	0.792	0.851
0.05	0.06	0.644	0.711
0.05	0.07	0.487	0.547

2.5 Interim Analyses

2.5.1 Early stopping for inferiority of treatment shortening arm

Interim analyses will be performed for safety in Arms B and C, to evaluate whether the poor outcome rate is worse in the arm with earlier treatment completion. The Kaplan-Meier based estimator of poor outcomes at week 72 will evaluated after 1/3 and 2/3 of participants have been followed for 72 weeks from study entry, using a spending function that mimics the Pocock boundary. The stopping boundary is derived from a test of inferiority (of the treatment shortening arm) that corresponds approximately to a z-score of 2.178 (i.e., a two-sided p-value of 0.029), although the precise boundary depends on the exact amount of information at the interim analyses.

2.5.2 Early stopping for study futility

The premise of this study is that imaging and Xpert markers can identify a subset of participants with high success rates. If the relapse rate is high in this subset, then the basis for this study must be called into question. Therefore, when about half of the participants have completed their week 72 follow-up, the poor outcome proportions will be evaluated and presented to the Data and Safety Monitoring Board (DSMB). If more than 16 (of 75) participants in the standard treatment arm (Arm B) have a poor outcome, a recommendation to stop the trial will be considered. Alternatively, a recommendation to stop randomization into Arm C will also be considered. In this scenario, participants who would be eligible for shortening would be put into Arm B. Under this scenario the lowest achievable poor outcome in the standard treatment arm (by the end of the trial) would be 11% (16/150). Poor outcome rates higher than 11% would be concerning given the eligibility criteria for randomization, which represents a subset of participants with a low probability of relapse. Conditional power will be included in the DSMB reports to give guidance about the likelihood of concluding non-inferiority if the study continues to full enrollment. A table of conditional power computations will be included using a range of

non-inferiority margins (e.g., ranging from 6%-10%), since determination of an acceptable margin may depend on multiple factors.

3 Analysis Populations

3.1 Modified intention-to-treat (mITT)

This study is different from many others in that participants must have adequate adherence for the 16 weeks of treatment prior to randomization. All randomized participants will be included in the mITT cohort, unless it is later determined that a participant is considered a protocol violation prior to randomization (e.g., documentated error in the PredictTB criteria). The following four scenarios will result in censoring at the times specified below.

1) Participant infected with a new strain (i.e., a TB strain that differs from that identified at baseline) after converting to culture negative. Participant will be censored at time of re-infection.
2) Participant died from a violent cause or trauma (e.g., traffic accident). Participant will be censored at time of violent death.
3) Participants who are lost to follow-up after completion of 6-month treatment phase with a negative last available LJ culture (for both arms B & C). Participants will be censored at the time of the last observed culture.
4) Participants who are able to produce sputum at week 72, but whose week 72 sputum samples are both contaminated or missing, who cannot be brought back for repeat cultures, provided they have not already been classified as unfavorable and provided that their last positive culture was followed by at least two negative cultures.

Note that participants lost to follow-up after week 16 and before week 24 will be considered as a poor outcome.

3.2 Per protocol

Additional participants will be excluded from the mITT subset for a per-protocol analysis. These include:

1) Participants who did not complete the treatment course to which they were randomized (e.g., a participant randomized to continue treatment until week 24, who stops treatment at week 20).
2) Participants whose treatment was modified or extended for reasons other than an unfavorable response to treatment.

3.3 Pharmacokinetic sub-studies

The Pharmacokinetic (PK) sub-study aims to identify those patients at highest risk of relapse to see if differences in sub-breakpoint MIC and/or PK/sub-breakpoint MIC were present at baseline. This will be compared to those who did not relapse. Based on preliminary data, we believe that participants who enter Arm A due to a poor response on the Week 4 PET/CT scan are at the highest risk of relapse. Hence, we will target this group of participants for inclusion into this PK sub-study. At week 4 the PET/CT scan from the main study will be collected. A matching participant (i.e., a "control") from the low risk arm will be similarly invited to participate. If consented and enrolled, participants will undergo PK sampling at week 12 and 16. The sub-study concludes after two days of PK sampling are

completed but no later than the week 24 study visit. Control participants who are later randomized to Arm C must complete PK sampling no later than the week 16 study visit as they will conclude treatment at that time.

4 Outcome Measures

Please refer to Figure 2 and Table 4 for the flow of outcomes. The primary efficacy endpoint will be the proportion of poor outcomes in arms B and C 72 weeks following treatment initiation. Poor outcomes include those that experienced a treatment failure (section 4.1), confirmed relapse (section 4.2), or died (non-violent death).

Solid culture results will be used for all primary endpoint analyses, with one exception. If the week 72 solid culture result is contaminated or missing and the participant cannot be brought back for an additional sputum sample, the liquid culture result may be used at the final week 72 study visit. For other study visits, only solid culture results will be used for the primary analysis. Solid culture results that are missing or contaminated will be classified as unavailable. Liquid culture results may be used for secondary analyses.

Participants randomized to Arms B or C who are subsequently found to have a positive culture for *Mtb* on solid medium between and including weeks 16-24 that is confirmed on a subsequent culture will be considered treatment failures. These participants will be referred to continue treatment per local SOC and followed observationally until the end of their treatment to determine outcome. TB DNA strain typing may be done (sec 5.5.3) to confirm whether this is the same strain of DNA as the participant had at baseline. Single positive cultures that are not confirmed on a subsequent sputum sample are not considered failures as these may have arisen from clerical error or laboratory contamination [1]

4.1 Treatment Failure

Participants who remain culture positive on solid medium at Week 24 in Arm A will be considered treatment failures and will be withdrawn from the study and referred to continue treatment per the local SOC. Participants who convert to solid culture negative and subsequently have a single solid culture positive for *Mtb* before or at week 24 need to have a subsequent culture positive for *Mtb* to be confirmed as treatment failures.

4.2 Treatment Recurrence and Relapse

Participants who convert their sputum to culture negative (2 consecutive negatives over ≥4 weeks) and who subsequently become culture positive for *M.tb* again on solid medium, during follow-up after week 24, confirmed by a second (on another day) sputum culture positive for *M.tb*, will be considered recurrences. Single positive cultures that are negative on follow-up culture will not be considered recurrences. Participants with a positive, contaminated, or unevaluable culture on the final month 18 (week 78) follow-up visit may be asked to return for sputum culture confirmation.

Relapses will be distinguished from re-infections by DNA strain typing and only relapses will be considered a study endpoint.

4.3 Treatment Success

Treatment success will be defined as a participant with at least 2 consecutive negative cultures on solid medium over a span of at least 4 weeks, achieved before the end of therapy, with no subsequent confirmed positive cultures during follow-up.

4.4 Losses to follow-up

Participants who are lost to follow-up prior to week 72, will be included in the analysis up to and including the last time culture results were received. From that point forward, they will be considered censored for the purposes of analyses.

5 Data analysis details

5.1 Baseline tables

Table 3. Summary of baseline characteristics will be reported overall and by country as below. Site-level summaries can be included in appendices for the primary manuscript

		Arm A (n=)	Arm B (n=)	Arm C (n=)
Sex	Males N (%)			
Age	Mean (Range)			
Weight (kg)	Mean (SD)			
BMI	Mean (SD)			
# of participants with at least one previous TB episodes not within the past 3 years	N (%)			
Current smoker	N (%)			
Previous smoker—those who are not current smokers but report a history of smoking	N (%)			
Duration of smoking (years)	Mean (Range)			
Number of cavities (>2 mL)				
None	N (%)			
1	N (%)			
2	N (%)			
3	N (%)			
>=4	N (%)			
Total cavity volume (>2 mL)	Mean (SD)			
Xpert CT – Week 0	Mean (SD)			
Time to positivity on MGIT (days) – Week 0	Mean (SD)			

Race (tabulate by reported categories)				

5.2 CONSORT diagram

A flow diagram describing the reasons for exclusion will be included following the CONSORT criteria [2].

5.3 Statistical analysis of primary outcome

The primary analysis will be based on the proportion of poor outcomes (as defined in section 4) at 72 weeks between Arms B and C based on the MITT population. A two-sided 90% confidence interval for the difference of the proportions (p_C-p_B, where p represents the proportion of poor outcomes and the subscripts indicate the treatment arm) will be based on proportion of poor outcomes from the Kaplan-Meier curve at week 72 and the Greenwood variance formula [3]. The advantage of using the Kaplan-Meier estimator at week 72 is the inclusion of participants who contribute outcome data after randomization but are lost to follow-up before week 72. Differences below zero indicate worse outcomes for arm C. Therefore, if the lower bound of the confidence interval is greater than or equal to -0.07, the null hypothesis will be rejected, and non-inferiority of arm C will be concluded.

Table 4. Primary outcome breakdown.

Variable	Arm B (N=)	Arm C (N=)	Arm A (N=)	
Favorable outcome – no. (%)				
Patients with outcome				
Unable to produce sputum				
Unable to produce sputum at 72 weeks but culture negative status earlier				
Missing data on LJ culture at 72 weeks and MGIT negative				
Poor outcome – no. (%)				
Through week24 (or through treatment if extended in Arm A)				
Death presumable related to TB				
Treatment failure – Culture-confirmed				
Treatment failure – Not culture-confirmed				
Withdrawal of consent				
Lost to follow-up				
No completion of treatment				
Adverse event severe enough to stop the treatment				
PET/CT findings that required immediate procedures or treatment				
Possible unfavorable outcome				
After Week 24 (or after treatment if extended in Arm A)				
Death presumable related to TB				
Treatment failure – Culture-confirmed				

Treatment failure – Not culture-confirmed				
Treatment relapse				
Lost to follow-up				
Possible unfavorable outcome				

5.4 Secondary Analyses

1) As a sensitivity analysis, the primary analysis will be repeated but using the per-protocol cohort.
2) As another sensitivity analysis, the primary analysis will be repeated using MGIT results to define final outcomes.
3) The difference (and 90% confidence interval) in treatment success rates between a combined A+B Arm and a combined Arm A+C, with weighting to represent the population. Let p_A and p_{BC} denote the proportion of A and B/C participants, respectively. Let F^A, F^B and F^C be the Kaplan-Meier estimates of success rates at week 72. The overall success rates for each arm can be computed as $p_A F^A + p_{BC} F^B$ and $p_A F^A + p_{BC} F^C$.
4) Univariate and multivariate logistic regression models will be fitted to demographic, radiographic, microbiologic and immunologic variables with outcomes of treatment failure vs cures. Markers from images include: PET total glycolytic activity in regions of interest, total volume of hard CT lesions (-100 to 100 HU), total volume of soft CT lesions (-500 to -100 HU), and cavity air (volume of air in cavities). Endpoints relating to immunologic markers will be based on serum cytokine levels as described in section 1. Xpert MTB/RIF cycle threshold will be analyzed as a continuous variable in addition to the pre-specified thresholds. Analyses will also consider transformations such as delta cycle threshold. A training and test split-sample approach will be employed as a validation step.
5) Univariate and multivariate logistic regression models will be fitted to demographic, radiographic, microbiologic and immunologic variables with outcomes of relapses vs cures, excluding treatment failures. Markers from images include: PET total glycolytic activity in regions of interest, total volume of hard CT lesions (-100 to 100 HU), total volume of soft CT lesions (-500 to -100 HU), and cavity air (volume of air in cavities). Endpoints relating to immunologic markers will be based on serum cytokine levels as described in section 1. Xpert MTB/RIF cycle threshold will be analyzed as a continuous variable in addition to the pre-specified thresholds. Analyses will also consider transformations such as delta cycle threshold. A training and test split-sample approach will be employed as a validation step.
4) Based on the limited number of models selected under 2) and 3) above, area under the ROC curves (AUC_{ROC}) will be estimated to summarize diagnostic accuracy. Time-dependent ROC curves will be estimated. [4,5]
5) To compare the diagnostic accuracy of TTP and GeneXpert cycle threshold as a marker of outcomes, analyses of these variables at different time points will be conducted. Specifically, AUC_{ROC}'s will be compared across time points.
6) ROC curve analysis will evaluate the accuracy of PK parameters of sub-breakpoint MIC, AUC/sub-breakpoint MIC, Cmax/sub-breakpoint MIC comparing treatment failures vs cures (and relapses vs cures), along with 95% confidence intervals.

Figure 2 TB outcome flow

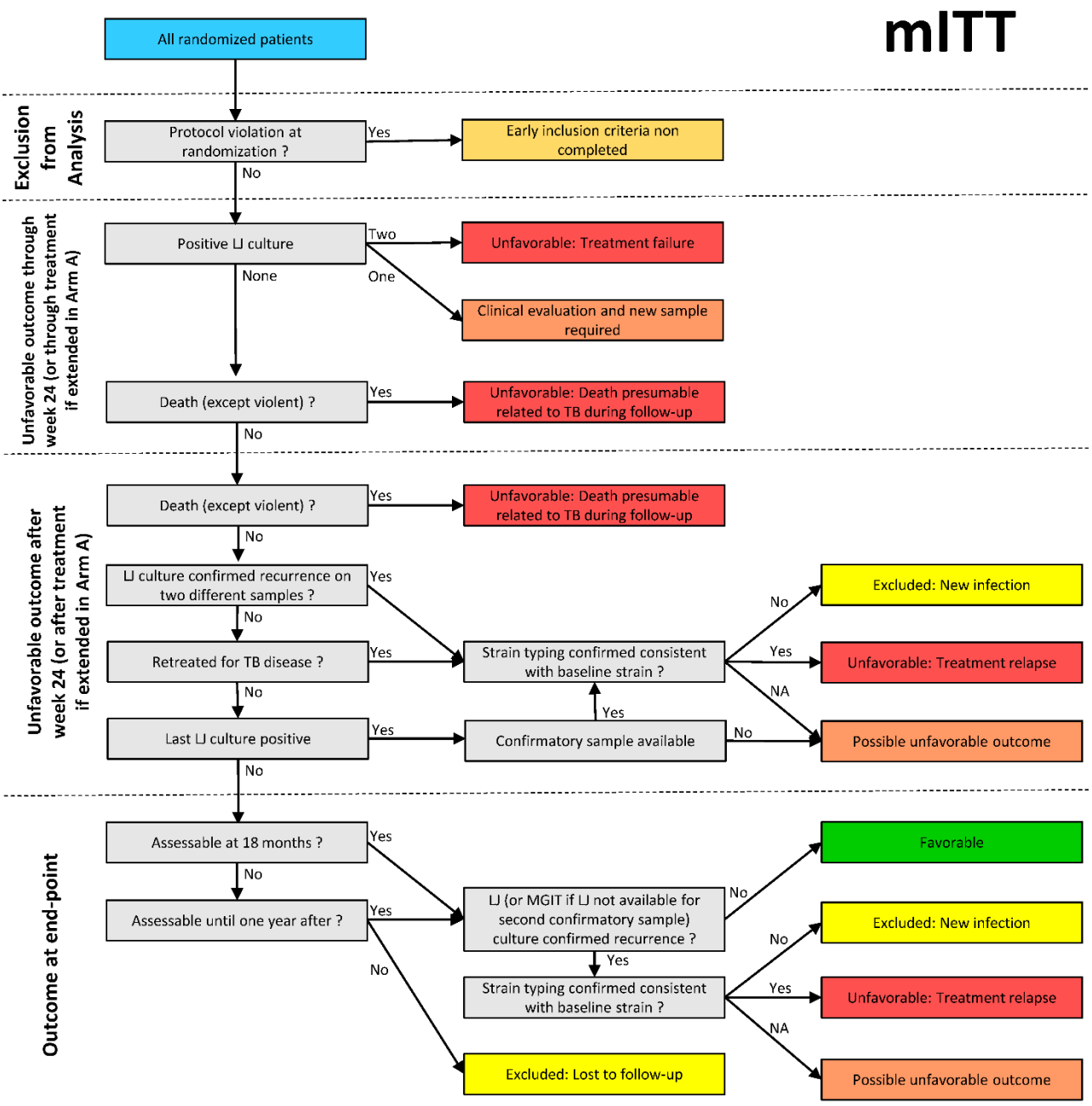

Table 5: TB Positive/Negative Flow for South Africa

LJ Culture:

CRF			Scenarios with LJ culture in SA (based on China)									
Test	Line	Item										
SPUTUM	6	Produced by	Induced ***or***	Spontaneous ***and***								
SPUTUM	7	Approx Volume	< 2 ml ***and***	>= 2 ml ***and***								
SPUTUM	8	Usage	Microbiology = Yes	Microbiology = Yes								
LJ	2	LJ Growth		NO	YES/Contaminated							
LJ	3	AFB SMEAR			ND/NA	AFB -	AFB + or Mixed					
LJ	4	Date		NA	NA	Date						
MGIT	8	TB antigen					Positive	Neg/Not done				
MGIT	9	Speciation result/HAIN						MTB complex or MTB/NTM	Any thing else			
LJ	5	TB antigen		Missing or Not done/Not available								
LJ	6	Speciation result/HAIN				Missing or Not required	Missing or Not required		MTB	NTM	MTB & NTM	Missing, Not r No res
RESULT			Alarm	TB -	Contaminated	Contaminated	TB +	TB +	TB +	TB -	TB +	Alarr
Outcome				TB-	Contaminated	Contaminated	TB +	TB+	TB+	TB-	TB+	Alarr

MGIT Culture:

CRF line			Scenarios with MGIT culture in RSA							
Test	Line	Item								
SPUTUM	6	Produced by	alarm	Spontaneous ***and***						
SPUTUM	7	Approx Volume	< 2 ml ***and***	>= 2 ml ***and***						
SPUTUM	8	Usage	Microbiology = Yes*	Microbiology = Yes*						
MGIT	3	MGIT result		Error / No result	Negative	Contaminated	Positive			
MGIT	4	TTD					DD:HH			
MGIT	5	AFB Smear					AFB + / Mixed			
MGIT	6	Date					Date			
MGIT	7	BAP					Pos / Neg / ND-NA			
MGIT	8	TB antigen					Positive	Negative		ND/NA/Missing
MGIT	9	Speciation result/HAIN						MTB complex or MTB/NTM	Any thing else	
RESULT			Alarm	Undetermined	TB -	Contaminated	TB +	TB +	TB -	Alarm
outcome				**Undetermined**	**TB -**	**Contaminated**	**TB +**	**TB +**	**TB -**	**Alarm**

Table 6: TB Positive/Negative Flow for China

LJ Culture:

CRF			Scenarios with LJ culture in China				
Test	Line	Item					
SPUTUM	6	Produced by	Induced ***or***	Spontaneous ***and***			
SPUTUM	7	Approx Volume	< 2 ml ***and***	>= 2 ml ***and***			
SPUTUM	8	Usage	Microbiology = Yes	Microbiology = Yes			
LJ	2	LJ Growth		NO	YES/Contaminated		
LJ	3	AFB SMEAR		Blank or NA	AFB -	AFB + / Mixed	ND/NA

LJ	4	Date		Blank or NA	Date	Date					Any value or missing
MGIT	8	TB antigen		do not reference	Any value or missing	Positive	Neg or Not done-NA **or blank**				Any value or missing
LJ	5	TB antigen		Blank or NA	Any value or missing	Any value or missing	Positive	Negative **or ND/NA**			Any value or missing
LJ	6	Speciation result/HAIN		Blank, Not required or No result	Any value or missing	Any value or missing	Any value or missing	Test not required, but if "MTB Complex" or "both MTB and NTM"	Test not required, but if "NTM" or "No Result"	Missing or Not required	Any value or missing
RESULT			Alarm	TB -	Contaminated	TB +	TB +	TB +	TB -	TB -	Alarm
Outcome			**edit check***	**TB-**	**Contaminated**	**TB +**	**TB +**	**TB +**	**TB -**	**TB -**	**Alarm**

* was able to get all these results with SPUTUM6= Induced or SPUTUM7 < 2ml

MGIT Culture:

CRF line			**Scenarios with MGIT culture in China**				
Test	**Line**	**Item**					
SPUTUM	6	Produced by	Induced ***or***	Spontaneous ***and***			
SPUTUM	7	Approx Volume	< 2 ml ***and***	>= 2 ml ***and***			
SPUTUM	8	Usage	Microbiology = Yes*	Microbiology = Yes*			
MGIT	3	MGIT result		Error / No result	Negative	Contaminated	Positive
MGIT	4	TTD		Blank or NA	Blank or NA	any	DD:HH
MGIT	5	AFB Smear		Blank or NA	Blank or NA	Blank or NA	AFB + / Mixed

MGIT	6	Date		Blank or NA	Blank or NA	Any value or missing	Date					
MGIT	7	BAP		Blank or NA	Blank or NA	Blank or NA	Pos / Neg / ND-NA			Blank or ND/NA	Pos/N	
MGIT	8	TB antigen		ND/NA	ND/NA	ND/NA	Positive	Negative	ND/NA/Missing	Blank or ND/NA	Pos	
RESULT			Alarm	Undetermined	TB -	Alarm	TB +	TB -	Alarm	Alarm	TB-	
		Outcome	**edit check***	**Undetermined**	**TB -**	**Contaminated**	**TB+**	**TB-**	**Alarm**	**Alarm**	**TB-**	

* was able to get all these results with SPUTUM6= Induced or SPUTUM7 < 2ml

References

1. Mitchison, D.A., et al., *Quality control in tuberculosis bacteriology. 2. The origin of isolated positive cultures from the sputum of patients in four studies of short course chemotherapy in Africa.* Tubercle, 1980. **61**(3): p. 135-44.
2. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMC medicine. 2010 Dec;8(1):18.
3. Greenwood M. The natural duration of cancer (report on public health and medical subjects no 33). London: Stationery Office. 1926.
4. Heagerty PJ, Zheng Y. Survival model predictive accuracy and ROC curves. Biometrics. 2005 Mar;61(1):92-105
5. Saha P, Heagerty PJ. Time-dependent predictive accuracy in the presence of competing risks. Biometrics. 2010 Dec;66(4):999-1011.