A Phase IIc, Open-Label, Randomized Controlled Trial of Ultra-Short Course Bedaquiline, Clofazimine, Pyrazinamide and Delamanid versus Standard Therapy for Drug-Susceptible Tuberculosis (PRESCIENT)

Primary Statistical Analysis Plan

Version 2.0

26 September 2024

ClinicalTrials.gov Identifier: NCT05556746

IND#: 164783

Protocol Version 5.0

(This is PRESCIENT SAP Version 2.0 with names of authors redacted)

Table of Contents

Version History

Version	Changes Made	Date Finalized
1.0	Original Version	October 26, 2023
1.1	Minor version update for protocol v4.0	April 30, 2024
2.0	Version update to align with protocol v5.0 Minor edits to correct typographic errors	September 26, 2024

1 Introduction

1.1 Purpose

This Primary Statistical Analysis Plan (SAP) describes the primary and key secondary estimands and other secondary outcome measures that will address specific study objectives and interim monitoring of the PRESCIENT (PRS01) study. The Primary SAP includes general analytic approaches for all primary estimands, key secondary estimands, and secondary outcome measures in the primary manuscript(s) or submitted to ClinicalTrials.gov (regardless of the reporting timeline).

2 Study Overview

2.1 Overview of Study Design

PRESCIENT is a Phase IIc, open-label, multi-center, randomized controlled trial to evaluate the microbiological efficacy of a rifamycin-free ultra-short regimen for drug-susceptible (DS) Tuberculosis (TB). The study population includes 156 adult (≥18 years) patients with confirmed smear-positive drug-susceptible pulmonary TB prior to receipt of > 5 days of TB treatment for the index episode, regardless of HIV status. Patients living with HIV are eligible if they have CD4 counts ≥200 cells/mm^3, if they are receiving ART or plan to initiate ART within 8 weeks after enrollment.

Eligible participants will be randomized in a 1:1 ratio to the experimental drug combination (Bedaquiline, Clofazimine, Pyrazinamide and Delamanid; BCZD for 12 weeks) or standard of care anti-TB therapy (Rifampicin, Isoniazid, Ethambutol, Pyrazinamide; RHZE for 26 weeks - 8 weeks of rifampicin, isoniazid, pyrazinamide and ethambutol (RHZE) followed by 18 weeks of rifampicin and isoniazid (RH)). Randomization will be stratified by presence of lung cavitation, HIV status and clinical site.

The primary analysis will be finalized once the last enrolled participant has completed the Week 10 study visit, culture data are available, and queries have been resolved. Secondary analyses will be finalized once the last enrolled participant has completed the Week 86 study visit, and all queries have been resolved.

2.2 Primary Hypothesis

Null hypothesis: No difference in time to culture conversion by 8 weeks between the experimental arm (Arm 1: BCZD) and standard therapy (Arm 2: RHZE).

Alternate hypothesis: The 12-week experimental regimen (Arm 1: BCZD) will demonstrate superior microbiologic efficacy (time to liquid culture conversion during the first 8 weeks of treatment) relative to standard therapy (Arm 2: RHZE).

2.3 Study Objectives

This Primary SAP addresses the following primary and secondary objectives in the study protocol. Analysis of the study objectives below will be analyzed under a superiority framework.

2.3.1 Primary Objective 1

To compare time to liquid culture conversion through 8 weeks for BCZD vs. standard treatment in HIV-positive and -negative adults with DS-TB.

2.3.2 Secondary Objective 1

To compare safety (adverse events) and tolerability over 60 weeks.

2.3.3 Secondary Objective 2

To compare cumulative probability of experiencing favorable composite outcome at 60 weeks between participants treated with BCZD versus with standard treatment.

2.3.4 Secondary Objective 3

To detect treatment-emergent resistance with BCZD.

2.3.5 Secondary Objective 4

To compare time to liquid culture conversion through 12 weeks for BCZD versus standard treatment.

2.3.6 Secondary Objective 5

To compare time (days) to positivity in liquid culture over 8 weeks for BCZD versus standard treatment.

2.4 Summary of Sample Size Considerations

This trial is powered for the primary outcome measure of comparing time to liquid culture conversion through 8 weeks in Arm 1 (experimental) versus Arm 2 (control). We estimate that 60% of Arm 2 participants will experience culture conversion within 8 weeks. We assume the true underlying hazard ratio to be 1.8 for Arm 1 vs. Arm 2, corresponding with a 35% relative increase in culture conversion over Arm 2. With a one-sided alpha level of 0.05 and 1:1 allocation to Arm 1: Arm 2, we would require 142 participants to have 90% power to reject the null hypothesis of no difference in time to culture conversion by 8 weeks between the two arms. To account for possible late exclusions, withdrawals, and missing culture results, we include a 10% inflation resulting in a total required sample size of about 156 participants. Additional details are provided in the trial protocol.

2.5 Overview of Formal Interim Monitoring

The study will be reviewed at least annually by a Data Safety Monitoring Board on a schedule outlined in the DSMB charter. For all reviews, the DSMB will be provided detailed information by randomization arm on safety, tolerability, and administrative aspects (including accrual, retention, and compliance with study requirements). In addition to routine monitoring of participant safety and study conduct, we will carefully monitor QT prolongation and hepatotoxicity. There will also be interim analyses monitoring early recurrence. Details of these analyses and triggers for DSMB review are described in the Study Progress Data and Safety Monitoring Plan (SPDSMP).

2.5.1 Early Assessment of QT Prolongation

There will be two assessments of QT prolongation. The first assessment will occur after 12 Arm 1 participants have evaluable ECG data available through week 2. We will find the point estimate of the proportion of participants on Arm 1 (who started assigned study treatment) with prolonged QT interval (defined by QTcF > 500 ms) and corresponding two-sided Clopper-Pearson exact 95% confidence interval. If the upper bound of the Clopper-Pearson exact 95% confidence interval around the proportion of Arm 1 participants exhibiting this AE exceeds a threshold of 40%, the team will evaluate the events carefully in conjunction with the DSMB to consider whether there should be a dose modification of QT-prolonging agents. A second assessment will occur similarly when the first 12 participants have evaluable ECG data available through week 12. There will be no adjustment to the confidence bands for multiple looks. Using data from 12 participants, if two or more participants experience prolonged QT, this evaluation will be prompted (i.e. the upper confidence limit would be >40%).

2.5.2 Early Assessment of Hepatotoxicity

There will be one early assessment of hepatotoxicity in Arm 1 participants who started assigned study treatment based on ALT and AST measured through week 12. This assessment will occur once 20 Arm 1 participants (who started assigned study treatment) have available AE data through week 12. We will find the point estimate of the proportion of participants with Grade 3 or higher hepatitis, defined as ALT or AST greater than 5X ULN, or clinically significant hepatitis (as defined in the protocol) and the corresponding two-sided Clopper-Pearson exact 95% confidence interval. If the upper bound of the Clopper-Pearson exact 95% confidence interval around the proportion of Arm 1 participants exhibiting this AE exceeds a threshold of 40%, the team will evaluate the events carefully in conjunction with the DSMB.

2.5.3 Early Recurrence Monitoring

TB recurrence monitoring will begin after the first Arm 1 participant completes 12 weeks of treatment and continue until the end of the study. Recurrence will capture confirmed relapses and reinfections. Since recurrence data on SOC will be lagging behind the experimental arm, recurrence data in experimental arms will be compared with historical SOC data from REMox,

OFLOTUB, RIFAQUIN, Study 31 (A5349) trials, and current SOC data within PRESCIENT as it becomes available. Two-sided 95% confidence bands for the cumulative probability of relapse over time will be generated monthly for the study team and ahead of each DSMB review. If the lower 95% confidence bound for the cumulative probability exceeds 3% at any time, this will trigger a DSMB review. There will be no adjustment to the confidence bands for multiple looks.

2.5.4 Early Unfavorable Outcome Monitoring

Monitoring of Unfavorable Composite TB Outcome Status in Arm 1 participants will begin after 25 participants have completed assigned study treatment or after 5 participants have experienced an unfavorable composite TB outcome, whichever comes first, and will continue approximately monthly. The DSMB will be notified if the point estimate of the proportion of Arm 1 participants with an unfavorable composite TB outcome exceeds 20%.

3 Outcome Measures

3.1 Primary Outcome Measure

3.1.1 Time to stable liquid culture conversion [Time Frame: Measured through Week 8]

Defined as the first of two negative sputum cultures, consecutive or not, without an intervening positive culture, and/or visits wherein the participant is unable to produce sputum and has no signs or symptoms of active TB.

3.2 Secondary Outcome Measures

3.2.1 Proportion experiencing any Grade 3 or higher AE including any occurrence that is new in onset or aggravated at least a one-grade from baseline [Time Frame: Measured at Week 60]

Graded according to the Division of AIDS Table for Grading the Severity of Adult and Pediatric Adverse Events (DAIDS AE Grading Table), Corrected Version 2.1, July 2017.

3.2.2 Proportion with favorable composite outcome [Time Frame: Measured at Week 60]

Additional details in protocol section 9.2.

3.2.3 Proportion who prematurely discontinue treatment [Time Frame: Measured at Week 12 in Arm 1 and Week 26 in Arm 2]

3.2.4 Time to stable liquid culture conversion [Time Frame: Measured through Week 12]

Defined as the first of two negative sputum cultures, consecutive or not, without an intervening positive culture, and/or visits wherein the participant is unable to produce sputum and has no signs or symptoms of active TB.

3.2.5 **Mean subjective 'change in coloration' score at weeks 8, 12, 16, 26, 60, and 86 in Arm 1 vs Arm 2 (10-point numeric rating scale). [Time Frame: Measured at Weeks 8, 12, 16, 26, 60, and 86]**

10-point numeric rating scale where 0=none, 10=worst possible change in coloration

3.2.6 **Mean subjective 'distress related to skin coloration' score at each of weeks 8, 12, 16, 26, 60, and 86 in Arm 1 vs Arm 2 (10-point numeric rating scale). [Time Frame: Measured at Weeks 8, 12, 16, 26, 60, and 86]**

10-point numeric rating scale where 0=none, 10=worst possible distress due to coloration

3.2.7 **Mean change in QTcF from baseline to week 1, 2, 4, 8, 12, and 16 [Time Frame: Measured at Weeks 1, 2, 4, 8, 12, and 16]**

The QTcF is derived from ECG readings, which the sites conduct in triplicate (three ECGs 5-10 minutes apart). The mean of all measurements (up to 3) that are readable and available will be used at each of baseline (screening visit), week 1, week 2, week4, week 8, week 12, and week 16.

3.2.8 **Mean change in QTcF from baseline to end of treatment [Time Frame: Measured at week 12 in Arm 1]**

The QTcF is derived from ECG readings, which the sites conduct in triplicate (three ECGs 5-10 minutes apart). The mean of all measurements (up to 3) that are readable and available will be used at baseline (screening visit), week 12 (Arm 1).

3.2.9 **Occurrence of absolute QTcF >480 ms and ≤500 ms, and >500 ms [Time Frame: Measured through week 16 in Arm 1 and Arm 2]**

The QTcF is derived from ECG readings, which the sites conduct in triplicate (three ECGs 5-10 minutes apart). The mean of all measurements (up to 3) that are readable and available will be used at week 16 (Arm 1 and 2).

3.2.10 **Occurrence of QTcF change from baseline of >30 ms and ≤60 ms, and >60 ms [Time Frame: Measured through week 16 in Arm 1 and in Arm 2]**

The QTcF is derived from ECG readings, which the sites conduct in triplicate (three ECGs 5-10 minutes apart). The mean of all measurements (up to 3) that are readable and available will be used at week 16 (Arm 1 and 2).

3.2.11 **Proportion of participants with one or more serious adverse events (SAEs) [Time Frame: Week 86]**

Serious adverse events reported at any time during participation in the trial.

3.2.12 **Proportion with culture conversion in liquid and solid media (separately) at weeks 4, 8, and 12 after randomization [Time Frame: Measured at Weeks 4, 8, and 12].**

Proportion of participants who have achieved stable culture conversion, defined as two negative sputum cultures, consecutive or not, without an intervening positive culture and/or visits wherein the participant is unable to produce sputum and has no signs or symptoms of active TB; occurring before or at the week 4, 8, or 12 visit, respectively.

3.2.13 **Proportion with TB relapse (by *M. tuberculosis* genotyping) from end of treatment to 86 weeks. [Time Frame: Measured at Week 86]**

For participants who had successful culture conversion through the end of study treatment, TB relapse is defined as a recurrence of TB emanating from the same strain as the participant's originally diagnosed TB, which will be determined through whole genome sequencing.

3.2.14 **Proportion of treatment-emergent genotypic and phenotypic resistance to BCZD. [Time Frame: Measured through Week 86]**

For participants in Arm 1 only. Minimum Inhibitory Concentration (MIC) values will be evaluated against resistance-associated variants (RAVs) for paired baseline and failure isolates. Frequencies and proportions with phenotypic and/or genotypic resistance to any drug will be reported.

3.2.15 **Time (days) to positivity in liquid culture (MGIT) after start of treatment across study arms [Time Frame: Measured through Week 8]**

Median (Q1, Q3) times to positivity in liquid culture at each time point (weeks 1, 2, 3, 4, 6, and 8) in Arm 1 and Arm 2. For visits where a participant has more than one available liquid culture result, the result with the positive finding will take precedence over a negative result.

4 General Considerations

All statistical comparisons for the outcome measures described in Section 3 are between Arms 1 and 2. There will be no statistical comparisons across groups for baseline characteristics. The analysis set is specified for the analysis of each study outcome measure.

Time-to-event outcomes will be analyzed using a Cox proportional hazard model. We will assess the proportional hazards assumption visually by plotting the scaled Schoenfeld residuals against the natural logarithm of the time to event. In the event that the data violate the proportional hazards assumption, we will conduct the primary analysis with a stratified log-rank test. Time to event outcomes may be supplemented with the difference in restricted mean survival times using the respective time horizon(s) defined in each outcome.

For analyses where we estimate cumulative probabilities with the Kaplan-Meier estimator at a fixed time point, we will use Greenwood's estimator to estimate the variance of the proportions in each arm, and the variance for the difference in proportions between arms will be computed as the sum of the squared variance estimates for Arms 1 and 2.

4.1 Analysis sets

- Efficacy Set:
  All participants who are randomized to study treatment and are not late exclusions. (Participants will be late exclusions if either of the following conditions are met: 1) Screening, Entry, and Week 1 visit sputum cultures (liquid or solid) all fail to grow M. tuberculosis. 2) Resistance to RIF or INH is detected from baseline (screening or entry visit) molecular or phenotypic testing results that become available after enrollment.

- Safety Set:
  All participants who start TB treatment.

5 Estimands and Estimation

5.1 Primary Estimand

<table>
<tr><td colspan="3">Primary Objective 1: To compare time to liquid culture conversion through 8 weeks for BCZD vs. standard treatment in HIV-positive and -negative adults with DS-TB.</td></tr>
<tr><td>Estimand description</td><td colspan="2">Time to stable liquid culture conversion through 8 weeks of follow-up among individuals aged ≥18 years and with pulmonary TB without demonstrated isoniazid or rifampicin resistance.</td></tr>
<tr><td>Treatment</td><td colspan="2">1) 12-week BCZD regimen
2) 26-week standard regimen</td></tr>
<tr><td>Target population</td><td>Analysis set</td></tr>
<tr><td>Individuals aged ≥18 years with confirmed pulmonary TB without demonstrated isoniazid or rifampicin resistance.</td><td>Efficacy set (all participants who are randomized to study treatment and are not late exclusions)
This analysis set will be further restricted to require that participants have a positive liquid culture at Entry, Screening, or Week 1.</td></tr>
<tr><td>Variable(s)</td><td>Outcome measure 3.1.1</td></tr>
<tr><td>Liquid TB culture results from serial collections of sputum from study visits up to week 10.

Note: We will consider a participant’s inability to produce sputum with no signs or symptoms of active TB at a given study visit as a negative sputum culture result.</td><td>Time to stable liquid culture conversion in liquid media, defined as the first of two negative sputum cultures, consecutive or not, without an intervening positive culture, and/or visits wherein the participant is unable to produce sputum and has no signs or symptoms of active TB, up to 8 weeks post-randomization.

(Note: We will also look at the liquid culture result from sputum collected at the Week 10 study visit if the first negative culture occurs at week 8.)</td></tr>
<tr><td>Handling of intercurrent events</td><td>Handling of missing data</td></tr>
<tr><td>The occurrence of any of the following intercurrent events through week 8 are relevant to the estimand:

1. Premature TB treatment discontinuation/treatment change (excluding death).
Observations will be used to determine the variable (treatment policy strategy)

2. TB-related death
Composite strategy (participant will be censored at week 8).

3. Non-TB-related death</td><td>Participants who have contaminated or missing samples leading up to and including week 8 will be censored at the last study visit for which they had a valid sputum culture assessment.

Participants who are lost to follow-up prior to 8 weeks with their last culture being positive will be censored at 8 weeks (i.e. assume they did not achieve culture conversion by week 8).

Participants who are lost to follow-up prior to 8 weeks with their last culture being negative will be censored at the</td></tr>
</table>

Composite strategy (participant will be censored at week 8).	last sampling date for which they had a valid culture result.
Population-level summary measures	Analysis approach
1. Primary: Hazard ratio of culture conversion through 8 weeks of follow-up. 2. Supplementary: Difference in restricted mean survival times over 8 weeks of follow-up.	We will estimate the treatment effect with a hazard ratio and corresponding 90% two-sided confidence interval. We consider Arm 1 superior to Arm 2 if the one-sided p-value (to test hypothesis that HR > 1) is less than 0.05 in favor of Arm 1. The hazard ratio (HR) will be estimated from a Cox proportional-hazards model adjusted for HIV status (positive; negative), lung cavitation and clinical site. In the event that the data violate the proportional hazards assumption, we will conduct the primary analysis with a stratified log-rank test.

Sensitivity Analyses

None.

Supplementary Analyses

We will estimate the difference in restricted mean survival times for a time horizon of 8 weeks and corresponding 90% CI. The RMST in each arm will be estimated by the area under the respective Kaplan-Meier curve.

We will also assess treatment by sex interaction terms in separate Cox proportional hazards models. These models will also include the same stratification factors as the primary analysis: HIV status, lung cavitation and clinical site.

5.2 Key Secondary Estimand (Safety)

<table>
<tr><th colspan="2">Secondary Objective 1: To compare the proportion of adverse events (safety) over 60 weeks between Arm 1 and Arm 2</th></tr>
<tr><td>Estimand description</td><td>Probability of having at least one grade 3 or higher adverse event through 60 weeks including any occurrence that is new in onset or aggravated at least one-grade from baseline among patients aged ≥18 years with pulmonary TB without demonstrated isoniazid or rifampicin resistance.</td></tr>
<tr><td>Treatment</td><td>1) 12-week BCZD regimen
2) 26-week standard regimen</td></tr>
<tr><td>Target population</td><td>Analysis set</td></tr>
<tr><td>Patients aged ≥18 years with pulmonary TB without demonstrated isoniazid or rifampicin resistance.</td><td>Safety Set (All participants who start TB treatment)</td></tr>
<tr><td>Variable(s)</td><td>Outcome measure 3.2.1</td></tr>
<tr><td>Occurrence of a Grade 3 or higher adverse event through 60 weeks that is new in onset or aggravated in severity of frequency from the baseline condition.</td><td>Proportion across study arms experiencing any Grade 3 or higher AE including any occurrence that is new in onset or aggravated at least a one-grade increase from baseline over 60 weeks.</td></tr>
<tr><td>Handling of intercurrent events</td><td>Handling of missing data</td></tr>
<tr><td>1. Persistent clinical disease and extension of treatment
Observations will be used to determine the variable (treatment policy strategy)

2. Premature treatment discontinuation
Observations will be used to determine the variable (treatment policy strategy)</td><td>Participants who discontinue follow-up before week 60 will have their outcome determined based on data available until the time of discontinuation (i.e., missing observations after discontinuation that would be relevant to the variable are considered missing at random).
Participants who discontinue follow-up before week 60 and have not yet experienced a grade 3 or higher AE will be censored at the time of their last study visit.

A sensitivity analysis will consider participants who discontinue follow-up before week 60 as having at least one Grade 3 or higher adverse event that is at least a one-grade increase from baseline.</td></tr>
<tr><td>Population-level summary measure</td><td>Analysis approach</td></tr>
<tr><td>Difference in cumulative probability of having at least one Grade 3 or higher AE that is new in onset or aggravated at least one-grade increase from baseline at any</td><td>We will compare the cumulative probability of participants experiencing a Grade 3 or higher AE that is new in onset or aggravated at least one-grade increase from baseline at any time</td></tr>
</table>

time during the 60 weeks after treatment initiation.	during the 60-week study period between Arm 1 and Arm 2 using difference in the Kaplan-Meier estimator at week 60 for each of the study arms. We will present the point estimate and 90% two-sided confidence intervals for the difference in cumulative probabilities.

Supplemental Analyses:

As supplementary analyses, we will also estimate the Kaplan-Meier proportions and confidence intervals for the outcome measure up to week 12, and up to week 26. These time horizons correspond to the end of treatment for Arm 1, and Arm 2, respectively.

5.3 Key Secondary Estimand (Favorable composite outcome)

Estimand Description: Difference in proportion with favorable outcomes 60 weeks after initiation of TB therapy among individuals aged ≥18 years and with pulmonary TB without demonstrated baseline isoniazid or rifampicin resistance.

We will define favorable composite outcome as the following (copied from protocol Section 9.2):

Favorable Outcome *(*any one of the following scenarios)

- Participants with liquid culture negative status at week 60
- Participants who are unable to produce a sputum specimen at week 60 and are without signs or symptoms of ongoing active TB
- Participants who produce a sputum specimen that is contaminated in two liquid cultures at week 60 but who are without symptoms/signs of ongoing active TB

For any of the above to be considered favorable, participants must not already have been considered unfavorable or unevaluable, as defined below.

Unfavorable Outcome

- Absence of cure
  A participant will be considered to have absence of bacteriological cure if a sputum sample obtained at or after end of treatment (week 12 for Arm 1;week 26 for Arm 2), that is culture-positive in liquid or solid media for an *Mtb* strain genotypically matched with the initial isolate. A second sputum sample, obtained on a separate sputum sample following the first sputum collection, is required to confirm absence of bacteriological cure. A second confirmatory sample is required as a single positive sputum culture in isolation will not be considered absence of bacteriological cure.

- Death
  Participants who die from any cause during study treatment or follow-up except from violent or accidental cause (e.g., road traffic accident).

- Last available culture positive

Participants who had a positive culture for *Mtb* when last seen, whether confirmed by a second sample or not.

- Treatment extension
  Participants experiencing extension of treatment beyond the nominal level (week 12 for Arm 1 and week 26 for Arm 2) due to clinically inadequate response. These cases should be managed individually in conjunction with and after having notified the Clinical Management Committee (CMC). Extension of treatment to make up missed doses will not count as unfavorable.

Unevaluable Outcome

- Participants lost to follow-up during treatment or follow-up with their last culture being negative for *Mtb*.
- Participants who die from violent or accidental death.
- Participants with recurrent TB due to a new strain of *Mtb* confirmed by whole genome sequencing.
- Women who become pregnant during their assigned active treatment and stop their assigned treatment.

Estimand specification:

Analysis Set: Efficacy Set: all participants who start their assigned study treatment and are not a late exclusion

Variable: Favorable outcome 60 weeks after initiation of TB therapy following evidence of cure at treatment completion where favorable outcome is defined above.

Intercurrent events:

- Minor treatment changes *(Treatment policy)*
- Participant has an isolated positive culture, but the definition of recurrence/absence of cure is not met, and culture conversion to stable negative status has been achieved. *(Treatment policy)*
- The participant is unable to produce sputum at the end of the follow-up period, having sustained culture negativity at the time the last sputum culture was obtained. *(Composite; favorable)*
- Death during treatment or follow-up except from violent or accidental cause. *(Composite strategy; unfavorable)*
- Death during treatment or follow-up from violent or accidental cause. (*"While on Treatment" strategy; restricts the observation time of interest to that before occurrence of the ICE*)

Missing data:

Participants lost to follow-up or otherwise unevaluable for the week 60 outcome will have their outcome imputed as follows:

- Participants with a positive culture for Mtb when last seen, whether confirmed by a second sample or not will be considered to have an unfavorable outcome.
- Participants lost to follow-up and culture negative at their last visit will be censored at the time of their last follow-up visit.
- Participants with a new TB diagnosis due to a new strain of Mtb will be censored at the time of their last follow-up visit prior to TB reinfection.
- Participants who become pregnant during assigned active treatment phase and who stop treatment will be censored at the time of treatment discontinuation.

- Population-level summary measures:
  - Difference in cumulative probability of having a favorable clinical/bacteriologic outcome at 60 weeks after TB therapy initiation.
  - We will estimate the difference in cumulative proportion of participants at week 60 together with 90% and 95% two-sided confidence intervals. We will use the Kaplan-Meier estimator and Greenwood's estimate for the standard error.

6 Analysis of Other Secondary Objectives

6.1 Premature treatment discontinuation (Objective 2.3.2, Outcome Measure 3.2.3)

- Analysis Set: Safety Set (all participants who start TB study treatment)
- Variable: Treatment discontinuation
- Intercurrent events:
  - Death during the treatment phase due to violent causes or natural disaster *(principal stratum strategy; participants who die from violent causes or natural disaster will be excluded from analysis)*
  - Death due to non-violent/accidental causes during treatment phase *(composite strategy; participants will be considered to have experienced premature treatment discontinuation)*

- Missing data: Those lost to follow up during the treatment phase will be assumed to have premature treatment discontinuation.

- Population-level summary measures:
  - Probability of permanent or premature discontinuation of study treatment

We will compare the proportion of participants who permanently and prematurely discontinue study treatment for any reason between Arm 1 and Arm 2. Permanent and premature discontinuation of study treatment is defined as not having completed an adequate course of

treatment (defined in protocol). The difference in proportions will be estimated using binomial proportions, and we will provide exact 95% confidence intervals.

6.2 Hyperpigmentation (Outcome Measures 3.2.5, 3.2.6)

- Analysis Set: Safety Set (all participants who start TB treatment)

- Variable: Skin coloration scores at Weeks 8, 12, 16, 26, 60, and 86.

- Population-level summary measures:
    - Mean ‘change in coloration’ score at each of weeks 8, 12, 16, 26, 60, and 86.
    - Mean ‘distress related to skin coloration’ score at each of weeks 8, 12, 16, 26, 60, and 86.

- Missing Data:
    - We will use only available data
    - In a sensitivity analysis, we will impute missing values using multiple imputation.

- Analysis:
  We will compare the ‘mean change in coloration score’ and ‘mean distress related to skin coloration score’ for Arm 1 vs Arm 2 at each of weeks 8, 12, 16, 26, 60, and 86 using Student T-tests. We will report the difference in means estimate and corresponding 95% confidence interval for each time point.

6.3 QT Prolongation (Outcome Measures 3.2.7, 3.2.8, 3.2.9, 3.2.10)

- Analysis Set: Safety Set (all participants who start TB treatment)

- Population-level summary measures:
    - Mean change in QTcF from baseline to weeks 1, 2, 4, 8, 12, and 16 in Arm 1 and Arm 2 (mean difference between arms for each change)
    - Mean change in QTcF from baseline to end of treatment (week 12 in Arm 1 and week 26 in Arm 2) (mean difference between arms for each change)
    - Odds ratio of occurrence of prolonged absolute QT interval (measured through week 16 in Arm 1 and 2)
    - Odds ratio of occurrence of prolonged change in QT interval (measured through week 16 in Arm 1 and 2)

We will compare the mean baseline QTcF in Arm 1 and Arm 2 with a two sample Student T-test and report the point estimate for the mean difference and corresponding 95% confidence interval. We will compare the QTcF in Arm 1 and Arm 2 similarly for time points of weeks 1, 2, 4, 8, 12, and 16.

We will compare the mean change in QTcF from baseline to weeks 1, 2, 4, 8, 12, and 16 between Arm 1 and Arm 2 using a GEE model and time by treatment interaction terms (treating time as a categorical variable). We will compare the change from baseline in QTcF over time between arms using a linear mixed effects model. Randomization arm and week (time) will be fixed effects and participant ID will be a random effect. In a supplemental analysis we will treat week (time) as a random effect. We will use the same methods to analyze the change in QTcF from baseline to end of treatment (Week 16 in Arm 1 and 2).

We will compare the odds of participants having maximum occurrence of absolute QTcF > 480 ms and ≤500 ms, or > 500 ms between Arm 1 and Arm 2 using a logistic regression model. We will compare the odds of participants having a QTcF change from baseline of > 30 ms and ≤60 ms, and > 60 ms, at any time during study treatment using logistic regression models.

6.4 Serious Adverse Events (Objective 2.3.2, Outcome Measure 3.2.11)

- Analysis Set: Safety Set (all participants who start TB treatment)

- Missing data: Participants who discontinue follow-up before week 86 will have their outcome determined based on data available until the time of discontinuation (i.e., missing observations after discontinuation that would be relevant to the variable are considered missing completely at random).

- Population-level summary measures: Probability of having at least one SAE of any grade at any time during the 86 weeks after treatment initiation.

- Sensitivity analyses: We will consider participants who discontinue follow-up before week 86 as having at least one SAE.

We will compare the cumulative proportion of participants with one or more SAEs of any grade between Arm 1 and Arm 2 using the Kaplan-Meier estimator at week 86. We will compute a participants' follow-up time (event or censoring) based on the actual date of their event or last study visit. If the last study visit occurs at week 86 within the 7-day visit window, we will mark the follow-up time as exactly 86 weeks post-randomization. We will present the point estimate for the difference in cumulative proportions and corresponding 90% two-sided confidence interval.

6.5 Culture conversion at weeks 4, 8, and 12 (Objective 2.3.1, Outcome Measure 3.2.12)

- Analysis Set: Efficacy Set (all participants who are randomized to study treatment and are not late exclusions)

- Missing data: Missing cultures at weeks 4, 8 and 12 will be imputed by the last observation carried forward.

- Population-level summary measures:
    - Probability of having achieved stable culture conversion at week 4 in solid media
    - Probability of having achieved stable culture conversion at week 4 in liquid media
    - Probability of having achieved stable culture conversion at week 8 in solid media
    - Probability of having achieved stable culture conversion at week 8 in liquid media
    - Probability of having achieved stable culture conversion at week 12 in solid media
    - Probability of having achieved stable culture conversion at week 12 in liquid media

- Sensitivity analyses: We will consider those with missing cultures to have not achieved stable culture conversion at the respective analysis week.

We will compare the binary proportion of participants with culture conversion at weeks 4, 8, and 12 between Arm 1 and Arm 2. We will analyze liquid and solid cultures separately. We will present the difference in proportions and corresponding exact 95% confidence interval for each comparison.

6.6 Relapse (Objective 2.3.4, Outcome Measure 3.2.13)

- Analysis Set: Efficacy Set (all participants who are randomized to study treatment and are not late exclusions)

- Missing data: Participants who are lost to follow up will be censored at the time of the last observed study visit.

- Population-level summary measures:
    - Probability of having TB relapse during the 86 weeks of follow up

- Sensitivity analyses: We will assume participants lost to follow up before 86 weeks have relapse.

We will estimate the cumulative probability of TB relapse for participants in Arm 1 and Arm 2 using Kaplan-Meier estimates and 95% confidence intervals at 86 weeks. We define the time of relapse as the time from end of treatment (after successful liquid culture conversion) until the first of two sputum samples that is culture positive in liquid or solid media for an *Mtb* strain that has matching genotype with the baseline isolate. A second positive culture obtained on a separate sputum sample after the first sputum collection is required to confirm a relapse. If the second sputum sample is negative, this will not be counted as a relapse. If a participant is lost to

follow-up without a second sputum sample confirmation of relapse (including contaminated or missing), it will be counted as a relapse.

For monitoring and interim reviews, we will estimate the proportion of participants with a recurrence (which captures both relapse and reinfection) over time with 95% pointwise confidence bands. This is because whole genome sequencing results will not be available until the end of the study.

6.7 Resistance (Objective 2.3.4, Outcome Measure 3.2.14)

- Analysis Set: Efficacy Set (all participants who are randomized to study treatment and are not late exclusions)

- Population-level summary measures:
    - Probability of having treatment-emergent phenotypical and genotypic resistance to BCZD during the 86 weeks of follow up

- Sensitivity analyses: We will assume those lost to follow up before 86 weeks developed resistance.

  We will plot the distribution of baseline MIC for each drug and describe frequencies of pre-treatment phenotypic resistance; MIC data will be incorporated into a pharmacodynamic model of TTP to explore influence on bacterial decline (treatment response), described below. MIC values will be evaluated against RAVs for paired baseline and failure isolates. We will cross-tabulate BDQ and CFZ MICs and a Pearson's correlation coefficient determined to assess cross resistance.

6.8 Time to positivity (Objective 2.3.6, Outcome Measure 3.2.15)

- Analysis Set: Efficacy Set (all participants who are randomized to study treatment and are not late exclusions)

- Population-level summary measures: Median (Q1, Q3) times (days) to positivity in liquid culture.

  We will summarize time (days) to positivity of TB cultures between Arm 1 and Arm 2 separately at weeks 1, 2, 3, 4, 6, and 8.

7 Primary Analysis Report Contents

- CONSORT diagram
- Study entry: Screening/Accrual/Eligibility
- Baseline characteristics (reported in ClinicalTrials.gov)
- Study retention (reported in ClinicalTrials.gov)
- Treatment status
    - Premature study treatment discontinuation reasons
    - Tolerance summaries
- Adverse events (reported in ClinicalTrials.gov)
- Analysis of primary outcome measures (reported in ClinicalTrials.gov)
- Analysis of secondary outcome measures (reported in ClinicalTrials.gov)
- Additional analyses as specified above