Approvals

IMPAACT P1108

Primary Statistical Analysis Plan

Version 2.1

A PHASE I/II, OPEN-LABEL, SINGLE ARM STUDY TO EVALUATE THE PHARMACOKINETICS, SAFETY AND TOLERABILITY OF BEDAQUILINE (BDQ) GIVEN IN COMBINATION WITH AN INDIVIDUALIZED RIFAMPIN-RESISTANT TUBERCULOSIS (RR-TB) THERAPY IN INFANTS, CHILDREN, AND ADOLESCENTS WITH RR-TB DISEASE, LIVING WITH OR WITHOUT HIV

Protocol Version 2.0, dated 21 September 2022
Letter of Amendment #1

ClinicalTrials.gov Identifier: NCT02906007

June 8, 2023

Created by:

Grace Montepiedra, PhD

Paula Britto, MS

SDAC/Harvard T.H. Chan School of Public Health

Table of Contents

Version History

Version	Changes Made	Date Finalized
1	Original Version	1/4/2017
1.1	Updates as per Letter of Amendments (LoAs) #2 and #3 and Clarification Memorandum (CM) #6 for protocol Version 1.0. To make it compliant with SDAC SOP requirements	05/27/2020
1.2	Version change due to LoA #4.	1/20/2021
2.0	Key changes 1. General alignment throughout SAP as per protocol V2.0 updates 2. Added NOTE to Section 3.2 to reflect updated serum creatinine grading requirements as per protocol Version 2.0 3. Added HIV-1 RNA and CD4 summary statistics, data Completeness, and Protocol deviations Summary to SMC reports in Section 4.2 4. Updated Section 4.3 “Safety Analysis Sets” to include all participants at the final “implemented dose” instead of final “recommended dose” per CM #1 for protocol Version 2.0.	04/25/2023
2.1	Minor revision due to LoA#1 for Protocol V2.0. No changes needed.	06/08/2023

1 Introduction

1.1 Purpose

This Primary Statistical Analysis Plan (SAP) describes the analysis plan for the primary and secondary outcome measures (except for the primary and secondary PK outcome measures) and additional outcome measures of the IMPAACT P1108 study that will be included in the primary manuscript(s). The Pharmacokinetic (PK) Statistical Analysis Plan (PK SAP) is a separate document prepared by the study pharmacometricians and will describe the analysis plan for the PK primary and secondary outcome measures. The Primary SAP outlines the general statistical approaches that will be used in the analysis of the study. It has been developed to facilitate discussion of the statistical analysis components among the protocol team, and to provide agreement between the protocol team and statisticians regarding the statistical analyses to be performed and presented in the primary analysis report. It also describes the results for the primary safety outcome measures and secondary outcomes measures (except for the PK secondary outcome measures) that will be posted on ClinicalTrials.gov.

Detailed outlines of tables, figures, and coding descriptions that will be included in the Analysis Reports are included in the Analysis Implementation Plan (AIP). There will be two Analysis Reports that will be prepared. The Primary Analysis Report will focus on the primary safety outcome measure and will be generated once the last participant has completed the Week 24 study visit and all data queries have been resolved. Analyses for the Secondary Analysis Report will address all secondary outcome measures, except for the PK outcome measures, and will be finalized once the last participant has completed the Week 96 study visit, all data queries have been resolved, and the study database closure/data lock has been completed. Outlines of analyses for exploratory objectives and outcome measures not included in the Primary SAP will be provided in a separate SAP.

1.2 Version History

The original version of the SAP was updated (V 1.1) to account for Letter of Amendment (LoA) #2, dated 28 February 2019, and LoA #3, dated 17 October 2019, for Protocol Version 1.0, dated 3 March 2016, and included updates to the Bedaquiline (BDQ) dosing regimen for all cohorts, as was determined by population PK modeling performed based on PK data for the first 12 evaluable participants in Cohort 1 (consistent with protocol Sections 3.2 and 10.2.3). In addition, this version also accounts for Clarification Memorandum (CM) #6, dated 1 April 2020, which provided guidance to sites for conducting study visits and procedures during the COVID-19 pandemic and recommended prioritization for conduct of clinically and scientifically important evaluations.

The next version 1.2 accounted for LoA # 4 dated 27 June 2022, which incorporated contents of protocol CM #6 that provided operational flexibility for conducting study visits and procedures when needed and to prioritize the conduct of clinically and scientifically important evaluations when possible. Consistent with the instructions provided in CM #6, implementation of this LoA is expected to be time-limited in relation to the COVID-19 pandemic.

Version 2 reflected updates based on Protocol Version 2.0, dated 21 September 2022, with CM #1, dated 29 March 2023.

This current version just notes the version change based on LoA #1 dated 2 May 2023 to Protocol Version 2.0/

2 Study Overview

2.1 Study Design

This is a Phase I/II, open-label, single arm study to evaluate the PK, safety, and tolerability of BDQ given in combination with an individualized rifampin-resistant tuberculosis (RR-TB) therapy in infants, children, and adolescents with RR-TB disease, living with or without HIV.

Sample Size: Up to 84 participants total (24 for Cohort 1; 30 for Cohorts 2; 30 for Cohort 3) to achieve 54 evaluable participants (18 per age cohort):

Cohort 1: Age ≥ 6 to < 18 years

Cohort 2: Age ≥ 2 to < 6 years

Cohort 3: Age ≥ 0 to < 2 years

Study Treatment and Duration: BDQ given as part of an individualized RR-TB treatment regimen for 24 weeks. For participants living with HIV, BDQ will be given in combination with an acceptable antiretroviral (ARV) therapy regimen initiated at least two weeks prior to enrollment.

<table>
<tr><th>Cohort</th><th>Age and Weight</th><th>BDQ Dosing</th></tr>
<tr><td rowspan="2">Cohort 1
up to 24 participants to achieve 18 evaluable (approximately nine in each weight band)</td><td>≥ 6 to < 18 years
≥30 kg</td><td rowspan="3">Participants ≥30 kg:
400 mg once per day through the intensive PK sampling visit*, then 200 mg three times per week on Monday, Wednesday, and Friday through the Week 24 visit

Participants >7 to <30 kg
200 mg once per day through the intensive PK sampling visit*, then 100 mg three times per week on Monday, Wednesday, and Friday through the Week 24 visit

Participants ≥ 3 to ≤ 7 kg:
100 mg once per day through the intensive PK sampling visit*, then 50 mg three times per week on Monday, Wednesday, and Friday through the Week 24 visit</td></tr>
<tr><td>≥ 6 to < 18 years
≥15 to <30 kg</td></tr>
<tr><td>Cohort 2
up to 30 participants to achieve 18 evaluable</td><td>≥ 2 to < 6 years
>7 kg</td></tr>
<tr><td>Cohort 3
up to 30 participants to achieve 18 evaluable</td><td>≥ 0 to < 2 years
≥3 kg</td><td></td></tr>
</table>

*Intensive PK sampling will be performed at the Week 1 or Week 2 visit following receipt of at least 14 and no more than 17 BDQ daily doses, including any non-study BDQ doses taken prior to study entry.

2.2 Study Objectives

2.2.1 Primary Objectives

- Determine the BDQ doses that achieve similar weekly exposure (area under the curve (AUC)) of BDQ compared to adults taking BDQ at the current standard recommended dose.
- Evaluate the safety and tolerability of BDQ over a 24-week dosing period.

2.2.2 Secondary Objectives

- Evaluate the PK of BDQ over the 24-week dosing period, by HIV status.
- Describe the long-term safety and tolerability of BDQ over a 96-week total follow-up period, by HIV status.
- Describe BDQ concentrations following BDQ discontinuation and through 72 weeks after BDQ discontinuation, by HIV status.
- Describe the RR-TB treatment response up to 96 weeks from initiation of the study, by HIV status.

2.3 Overview of Sample Size Considerations

At least 54 and up to 84 participants will be enrolled in the study across three age cohorts, as follows:

- Cohort 1: ≥ 6 to < 18 years of age at enrollment
- Cohort 2: ≥ 2 to < 6 years of age at enrollment
- Cohort 3: ≥ 0 to < 2 years of age at enrollment

Up to 24 participants in Cohort 1, up to 30 participants in Cohort 2, and up to 30 participants in Cohort 3 will be enrolled to achieve at least 18 evaluable participants in each cohort. In Cohort 1, up to 12 participants will be enrolled in each of the two weight bands (15 to < 30 kg and ≥30 kg) to achieve approximately nine evaluable participants in each weight band.

Participants are considered PK-evaluable if they have completed the intensive PK sampling collection and have at least one sample with measurable BDQ and M2 concentrations. Participants will be assessed by the Core Team for evaluability once PK data from the intensive PK visit are available. Regardless of evaluability, the safety population for this study will include all participants who receive at least one dose of study drug.

Non-evaluable participants will be replaced unless the maximum accrual of 24 participants in Cohort 1 and 30 participants in Cohorts 2 and 3 is already achieved. It is possible for the maximum number of participants to be enrolled in a cohort and not achieve the required 18 evaluable participants; however, the sample size per cohort was determined to accommodate loss of participant data due to non-evaluability or loss to follow-up.

The sample size is primarily based on PK considerations. Clinical trial simulations were performed to ensure a sample size able to provide precise enough estimates of apparent clearance as specified by the US Food and Drug Administration (FDA) criteria for pediatric trials. Since BDQ exposure over long-term administration primarily is determined by the apparent clearance, the first point implies that the primary objective of the study (to determine BDQ doses for children that achieve similar weekly exposure as adults taking BDQ at the current standard recommended dose) can be fulfilled.

2.4 Overview of Monitoring

2.4.1 Interim Monitoring

There will be two interim analyses of PK and safety data for each age cohort, for a particular dosing scheme, during the dose-finding phase. The first will occur when six participants under the current BDQ dosing scheme have completed their intensive PK sampling at the Week 1 or Week 2 visit (see Sections 6.3 and 6.4 of the protocol). If the first six participants have acceptable PK and safety profiles, as evaluated by the Core Team, a second interim analysis will occur when 12 participants under the current BDQ dosing scheme have completed their intensive PK sampling.

During each interim analysis, the safety data will be considered "unacceptable" if the following conditions are met, as assessed by the Core Team (*only among evaluable participants in the cohort under evaluation for the current dosing scheme*):

1) Any fatal or life-threatening adverse event (AE) assessed by the Core Team to be at least possibly related to study drug occurs.
2) For N=6: At least two participants experience a grade 3 or non-life-threatening grade 4 AE assessed by the Core Team as at least probably related to study drug.
3) For N=12: At least three participants experience a grade 3 or non-life-threatening grade 4 AE assessed by the Core Team to be at least probably related to study drug.
4) For N=18: At least four participants experience a grade 3 or non-life-threatening grade 4 AE assessed by the Core Team to be at least probably related to study drug.

The safety data will be reviewed concurrently with PK data and whether PK criteria are met or not.

NOTE: At the time of finalization of this document, all planned interim PK and safety analyses for all cohorts have been completed.

2.4.2 Monitoring by Study Monitoring Committee (SMC)

An IMPAACT SMC will review the study. The frequency of SMC review will be determined by the accrual rate or specific AEs and is planned to occur at least annually. Reports for the annual review may be minimal if accrual rate is slow. The SMC may also be convened upon request of the protocol team.

During the annual review, the Core Team may request the SMC focus on specific safety aspects, such as ECG data review. Note that SMC reports may be minimal if these types of events are rare and/or if accrual has been slow.

In addition, an *ad hoc* SMC review or consultation will occur, as noted below, by the following:

(1) In the event of **a fatal or life-threatening AE, as assessed by the site in consultation with the Core Team**, the Core Team will review the AE and assess its relationship to study drug.

- If the site investigator and/or the Core Team assesses the AE as **possibly, probably, or definitely related to study drug**, accrual will be paused. The Core Team will discuss how the study should proceed and consult with the SMC.
- If the site investigator and the Core Team assess the AE as **probably not or not related to study drug**, accrual will continue. The SMC will be informed of the AE along with the Core Team's assessment and decision-making.

(2) Within each cohort, and under the current dosing regimen: For N=6, if at least two participants within a cohort experience a grade 3 or non-life-threatening grade 4 AE at least probably related to study drug as assessed by the Core Team; OR for N=12, if at least three participants within a cohort experience a grade 3 or non-life-threatening grade 4 AE at least probably related to study drug as assessed by the Core Team; OR for N=18, if at least four participants within a cohort experience a grade 3 or non-life-threatening grade 4 AE at least probably related to study drug as assessed by the Core Team, then the SMC will be notified of this.

(3) If > 25% of participants within a given cohort experience a QTcF > 500 ms then the SMC will be notified of this.

(4) In the event of any unresolvable disagreement within the Core Team on an issue that would impact decision-making or if the Core Team encounters any other event or trend of concern, an SMC review of the relevant data will be convened.

3 Outcome Measures

3.1 Primary Outcome Measures (evaluated through week 24)

- Termination from treatment due to a drug-related AE
- AEs of ≥ grade 3 severity
- AEs of ≥ grade 3 severity assessed by the Core Team to be at least possibly related to the study drug
- Absolute QTcF interval ≥ 500msec
- Unstable dysrhythmias requiring hospitalization and treatment
- Death (grade 5 AE)

3.2 Secondary Outcome Measures (evaluated through Week 96 / 72 weeks post BDQ discontinuation)

- AEs of ≥ grade 3 severity

- AEs of ≥ grade 3 severity assessed by the Core Team to be at least possibly related to the study drug
- Absolute QTcF interval ≥ 500msec
- Unstable dysrhythmias requiring hospitalization and treatment
- Death (grade 5 AE)
  NOTE: For the outcome measures 3.1 and 3.2 above, for creatinine events under protocol Version 1.0, serum creatinine grading for the study analyses will be based on the absolute value only as per the DAIDS AE Grading Table (Corrected Version 2.1, dated July 2017) and not on percentage change from the participant's baseline value. This is consistent with creatinine grading requirements in protocol Version 2.0.

3.3 Secondary TB treatment outcomes

- TB treatment outcome (see Protocol Section **Error! Reference source not found.**) at Week 96/End of Study or Early Study D/C visit, classified as bacteriological cure with no TB recurrence, probable cure with no TB recurrence, TB recurrence, death, treatment failure, or lost to follow-up.
- TB treatment outcome (see Protocol Section **Error! Reference source not found.**) at end of RR-TB treatment, classified as bacteriological cure, probable cure, death, treatment failure, or lost to follow-up.

4 Statistical Principles

4.1 General Considerations

The primary safety analyses will focus on the 24-week time period during BDQ treatment and will include only participants whose total exposure to BDQ was at the final implemented dose for their cohort for the protocol-specified period of BDQ administration. Participants who have been removed from treatment, or who have had their doses reduced as part of cohort management due to toxicities, will be included and treated as safety failures in the primary safety analysis (note that such participants may have to be excluded from any secondary analyses which require complete follow-up at the optimal dose). Participants whose doses have been adjusted on the basis of PK results will be excluded from the primary safety analyses, regardless if the participant initiated BDQ at the final implemented dose, and sensitivity analyses performed in an attempt to determine whether the exclusion of these participants creates a selection bias which impacts upon any results. These primary analyses will be performed after the last participant of the last cohort has completed the study drug regimen over the 24-week dosing period.

Each participant's safety data will be summarized as: (1) the most severe grade of AEs, and (2) the most severe grade of AEs assessed to be at least possibly related to study treatment. Frequency distributions of these safety outcomes will be presented in aggregate and will be broken down by age cohort. Listings of all ≥ grade 3 events will be provided.

The proportions of participants experiencing ≥ grade 3 AEs will be presented in aggregate and broken down by age cohort, with these proportions bounded by exact 95% CIs. If sample size allows, frequencies and proportions will be presented by weight band; similarly, by HIV status, broken down by age cohort and/or in aggregate. Similar analyses will present the proportions of participants exhibiting ≥ grade 3 events which have been assessed to be at least possibly related to study drug, again bounded by exact 95% CIs. Tabulations will also be presented to summarize all AEs, as well as all AEs which have resulted in treatment discontinuation. Summary statistics of QT values at each time point performed will also be presented. Additionally, listings of participants who experience unstable dysrhythmias requiring hospitalization and treatment will be presented.

In addition, if possible, a primary evaluation of safety across the 24 weeks of study treatment will be performed on the data from participants who initiated the final recommended BDQ dose for a given cohort and remained on that dose for the 24-week period or have left the study or had a dose modification due to safety failure prior to 24 weeks of exposure (in which case the participant will be analyzed as a failure). Note that such an analysis may not be possible, since the PK modeling procedure which will determine the final recommended dose will not guarantee that any participant, or an adequate number of participants, be on that dose. However, secondary safety analyses will include all safety data collected from first participant exposure to the end of the study, with results broken down by dose. This will include data representing the final implemented dose for each cohort, as well as data gathered during the dose finding stage, which may represent exposure to doses which have failed.

Given that the modest sample sizes within cohorts will provide limited power for statistical tests of differences across age cohorts, only very large apparent effects would be statistically significant. Interpretation of differences across cohorts will depend upon whether these differences are great enough to be considered to be clinically significant. If no such differences are observed, then the clearest interpretation of the findings will come from the aggregated data, where analyses will have the greatest statistical precision. However, if results vary across cohorts to a clinically important extent, interpretation of results should consider the age differences and potential treatment differences represented by this stratification factor.

The proportions of participants meeting each of the endpoints which would trigger an *ad hoc* SMC review will be presented descriptively. Details concerning the analyses will be included in a separate analysis plan.

4.2 Analysis Approaches

This section contains the details of the analyses for:

- Interim and full cohort accrual safety reports [NOTE: At the time of finalization of this document, all planned interim PK and safety analyses for all cohorts have been completed]
- SMC Reports
- SDAC final reports

Summary tables will be generated by cohort and aggregated.

The interim and full cohort accrual safety reports will be prepared for the Core Team and will include Participant and Study Status and Safety Summary tables (these may be shared with the SMC).

The SMC reports will include information concerning Screening and Entry, Participant and Study Status, Baseline Characteristics, TB Treatment Outcome (if data are available), Safety summary tables, HIV-1 RNA and CD4 summary statistics, Data Completeness, and Protocol deviations Summary. Unless otherwise noted, the SDAC final reports will contain all summary tables.

Additional participant listings, not mentioned in this plan, but which may be necessary in writing up the SDAC final report, will be generated and will be used internally at SDAC. Highlights of these listings will be included in the SDAC final report. These participant listings will be provided to the manuscript writing team. The manuscript writing team, which is a subset of the protocol team, will be identified by the protocol team.

The analysis of the PK data will be performed by the protocol pharmacometricians and, thus, details of this analysis are not included in this document.

Validation requirements as per SDAC SOPs will be as follows:

a. All dataset creation programs will be validated with double coding for the primary and secondary safety endpoints.
b. All study-specific formats will be validated.
c. Validation of analysis programs is specified in the subsections below.

For all analyses, the relative weeks of visit date/specimen date/onset date/other dates (“Assessment Window”) will be calculated using the following guidelines:

NOTE: Protocol Appendix XII provides site guidance for conducting study visits and procedures during operational disruptions due to the COVID-19 pandemic.

Study Day of Assessment	Assessment Window	Target Study Day of Window
0 to 1	Week 0	0
2 to 10	Week 1	7
11 to 20	Week 2	13
21 to 35	Week 4	28

36 to 49	Week 6	42
50 to 69	Week 8	56
70 to 98	Week 12	84
99 to 126	Week 16	112
127 to 160	Week 20	140
161 to 195	Week 24	168
196 to 252	Week 32	224
253 to 308	Week 40	280
309 to 364	Week 48	336
365 to 462	Week 60	420
463 to 546	Week 72	504
547 to 714	Week 96/End of Study visit	672
(Study day of last dose + 14)	Early Discontinuation	Study day of last dose + 1

4.3 Safety Analysis Sets

The primary analysis set consists of all participants whose total exposure to the study drug BDQ was at the final implemented dose for their cohort for the protocol-specified period of BDQ administration. Participants who discontinued study drug due to toxicities will be considered a safety failure for this primary safety outcome measure* and will be considered an intercurrent event for the other primary safety outcome measures. Those who have had their doses adjusted as part of cohort management due to toxicities will be treated as safety failures* in the primary safety analysis. Participants who have been removed from study treatment for any reason (other than toxicity*) will be considered as having experienced an intercurrent event and will be included in the primary analysis dataset. These participants will be analyzed with the approach that those who experienced a primary safety outcome measure while on study drug will be considered as safety failure for that outcome measure. Those who were off study before the Week 24 visit due to reasons other than death will be included and will not be considered as a safety failure unless this was experienced before they were off study and while on study drug. The following sensitivity analyses will also be performed for participants who were off study before the Week 24 visit:

- The participant will be treated as safety failure (worst case)
- The participant will be excluded from the safety analysis set

Supplementary safety analysis dataset: All participants who were exposed to the study drug BDQ, including those who did not receive the final implemented dose evaluated in the study, as well as those whose dosing was amended based on PK considerations. The analysis approach will follow the same approach as that for the primary analysis set.

*Note that this refers to the relevant primary safety outcome (for example if the toxicity is “AEs of at least Grade 3 severity assessed by the Core Team to be at least possibly

related to study drug" then the participant will be considered a safety failure for this particular outcome measure only)

5 Report Contents

Detailed descriptions of the content of each of the following sections are given in the AIP.

1. Accrual
   - Screening (will be summarized in text only)
     - Number screened
     - Screening failure reasons
   - Enrollment
     - Number enrolled
2. Protocol deviations
   - Summary of reportable protocol deviations per IMPAACT Network Manual of Procedures
3. Baseline characteristics
4. Summary table for key characteristics, including sex, age, race, ethnicity, weight, height, Growth Z-scores, TB diagnostic category, TB disease type, and HIV status, TB Medication History, and key laboratory parameters
5. Study status
   - Number on treatment
   - Number off study prematurely, by reason and arm
   - Consort diagram (for final analyses only)
6. Safety Summary, in aggregate and by cohort
   - The proportions and listing of participants experiencing AE triggers as defined in protocol Section 9.5.2.
   - The listings and proportions of participants meeting the following primary endpoints:
     - Termination from treatment due to a drug-related AE
     - AEs of grade ≥ 3 severity
     - AEs of grade ≥ 3 severity assessed by the Core Team to be at least possibly related to the study drug
     - Absolute QTc interval ≥ 500msec
     - Unstable dysrhythmias requiring hospitalization and treatment
     - Death (grade 5 AE)
   - Listing of QTcF over time (including mean QTcF across all participants per study week)
7. TB Treatment Outcome (included in SMC report if data are available)

8. HIV-1 RNA and CD4 summary statistics, and ARV experience for participants living with HIV
9. Data Completeness Report (no tables, summary will be provided in text)

6 Associated Documents

Attachment 1: Writing Team Roster

Protocol Chair:	Anneke C. Hesseling
Protocol Vice Chair:	H. Simon Schaaf
NIAID Medical Officers:	Patrick Jean-Philippe
	Renee Browning
NICHD Medical Officers:	Sai Majji
	Jack Moye
Statisticians:	Grace Montepiedra
	Paula Britto
Protocol Pharmacometrician:	Elin Svensson
Protocol Data Managers:	Mattie Bartlett
	Amanda Golner
Clinical Research Managers:	Sarah Bradford
	Iris Mustich

Attachment 2: Timetable for primary analysis and manuscript preparation

Event	Responsible party	Weeks from primary completion date (PCD)
Primary completion date (PCD)		PCD
Clinical data entry complete	Sites	PCD + 2 weeks
Initial database clean complete	PDM and LDM	PCD + 13 weeks
Clinical database closure/freeze complete/eCRF sign-off by site investigators	PDM and Sites	PCD + 16 weeks
Clinical database lock and primary laboratory database lock	Chief Data Manager and LDM Leadership	PCD + 19 weeks
Primary analysis report to Protocol Chairs/Writing Team	Protocol statisticians	PCD + 27 weeks
ClinicalTrials.gov results for primary outcome measures due	Protocol statisticians	PCD + 52 weeks*

*The primary outcome measure results must be entered in clinicaltrials.gov exactly 1 year after the PCD.