IMPAACT P1078 Statistical Analysis Plan

Final Version 2.0 (Primary Analysis)

A Phase IV Randomized Double-Blind Placebo-Controlled Trial to Evaluate the Safety of Immediate (Antepartum-Initiated) Versus Deferred (Postpartum-Initiated) Isoniazid Preventive Therapy Among HIV-Infected Women in High TB Incidence

NCT 01494038

November 2, 2017

This is IMPAACT P1078 SAP Version 2.0 with names of authors, names of publication writing team members, and appendices redacted.

TABLE OF CONTENTS

Introduction

1.1 Purpose of the Statistical Analysis Plan (SAP)

This statistical analysis plan has been developed to (1) provide a structured time-frame for database completion and preparation of the final primary analysis report; (2) facilitate discussion of the key statistical analysis components amongst the study team; (3) provide agreement between the study team and statisticians regarding the statistical analyses to be performed; and (4) aid in setting priorities relating to the data cleaning and development of statistical analysis programs.

The statistical analysis plan contains details of all statistical analyses that will be completed for the final primary analysis report.

1.2 Study Design (Protocol Version 2.0)

This study is a randomized, double-blind, placebo-controlled non-inferiority trial designed to assess the overall safety and toxicity of 28 weeks of isoniazid (INH) preventive therapy (IPT) initiated during pregnancy versus deferred initiation of IPT at 12 weeks postpartum in HIV-infected pregnant women and their infants at high risk for TB infection and disease. The study will also assess the efficacy of these strategies in the women and their infants, safety and toxicity of INH *in utero* and on infants, INH resistance among Mycobacterium tuberculosis (*M.tb.*) isolates among participants who develop TB while on study, intensive PK of INH and ARV drugs among HIV-infected women on HAART, adherence to INH during pregnancy and postpartum, and performance characteristics of IGRA with TST among HIV-infected pregnant women and their infants.

A target of 950 HIV-infected pregnant women at ≥14 to ≤34 weeks (34 weeks, 6 days) of gestation and living in high TB incidence areas, defined as having 60 TB cases per 100,000 population in the WHO TB annual report, will be enrolled. Women who are suspected of having TB with a positive WHO TB symptom screen, have reported recent exposure to an active TB case, have received INH or TB treatment for >30 days in the past year, have evidence of acute hepatitis within 90 days prior to study entry, or have ≥Grade 1 peripheral neuropathy will not be eligible.

Participants will be stratified based on gestational age at entry, and will be randomized in equal proportions to one of the following two groups:

- ARM A: Maternal INH at study entry for 28 weeks and then switched to placebo for INH until 40 weeks postpartum (immediate IPT), or

- ARM B: Maternal placebo for INH from study entry until 12 weeks postpartum and then switched to INH for 28 weeks until 40 weeks postpartum (deferred IPT)

Each maternal participant will also receive vitamin B6 and multivitamins from study entry until 40 weeks postpartum, as well as intensified TB case finding, and will be followed up to 48 weeks postpartum. Each infant participant will be followed from birth until 48 weeks postpartum with

intensified TB case finding.

The primary hypothesis is that initiation of IPT during pregnancy among HIV-infected pregnant women is as safe as deferring IPT to 12 weeks postpartum.

2 Study Objectives (Protocol Version 2.0)

The primary analysis report will include the primary objective and secondary objectives 1-4 (listed in sections 2.1 and 2.2 below). Secondary objectives 5-7 and the exploratory objectives (section 2.3) will be addressed in separate analysis plans.

2.1 Primary Objective

To compare overall safety and toxicity of immediate versus deferred INH preventive therapy in HIV-infected pregnant women enrolled at ≥14 through ≤34 weeks gestation (34 weeks, 6 days)

2.2 Secondary Objectives

1. To compare safety and toxicity of INH *in utero* exposure and on infants on study
2. To compare TB incidence and all-cause mortality in HIV-infected women and their infants enrolled on study
3. To compare overall safety and hepatotoxicity, as well as evaluate risk factors for these outcomes, during pregnancy and immediate postpartum in women on immediate versus deferred INH therapy
4. To evaluate for INH resistance among *M.tb.* isolates from HIV-infected women and infants who develop TB while on study
5. To evaluate the intensive pharmacokinetics of INH and selected ARV drugs in a subset of HIV-infected pregnant and postpartum women receiving HAART
6. To evaluate and compare the performance characteristics of IGRA (TB ELISPOT and QGIT) with TST in HIV-infected women and their infants
7. To compare adherence in women initiating immediate versus deferred INH preventive therapy via self-report and pill counts

2.3 Exploratory Objectives

1. To evaluate population pharmacokinetics and pharmacogenomics of INH and EFV in HIV-infected pregnant and postpartum women
2. To assess the effects of INH preventive therapy on the functional characteristics of TB specific T-cell responses measured by TB ELISPOT in HIV-infected women and their infants on study
3. To evaluate adherence and exposure to INH and EFV using drug levels in hair and analyze the association between these drug levels and adverse events
4. To compare the two arms with respect to the overall (risk:benefit) clinical outcome of the mother-infant pairs
5. To explore associations of maternal and infant specific TB responses and novel biomarkers with risk of maternal and infant TB infection and disease

6. To explore the neurotoxicity of INH in combination with EFV in a subset of women

A brief description of the important study time points is shown in the following table:

Evaluations	Time points of interest
Adverse events in mothers (for primary safety analysis)	Up to week 48 postpartum
Adverse events in infants	Up to week 48 from birth
Adverse pregnancy outcomes	Delivery to week 48 from birth
Maternal TB	Up to week 48 postpartum
Infant TB	Up to week 48 from birth
Intensive and sparse PK in mothers	Antepartum (third trimester) and week 16 postpartum
TST in mothers	Delivery (or Week 4 postpartum if not done at delivery) and Week 44 postpartum
TST in infants	Week 44 from birth
IGRA in mothers	Entry, delivery and Week 44 postpartum
IGRA in infants	Week 44 from birth

3 Outcome Measures

3.1 Primary Endpoint

Grade ≥ 3 AEs possibly, probably, or definitely associated with INH/Placebo for INH or permanent discontinuation of INH/Placebo for INH due to an adverse reaction in women after randomization until 48 weeks postpartum.

Study drug attribution is based on adjudication by an Independent Endpoint Review Committee (IERC), which consists of OB/GYN and pediatric specialists who are not members of the study team. Refer to the Endpoint Review Process document for details of the review process.

Adverse events include laboratory results, signs/symptoms, or diagnoses reported in the study database. Note that for hematology results, only those collected at delivery are considered in the primary endpoint since testing at other study visits specified in protocol Version 1.0 was discontinued in Version 2.0. Except for the following, diagnoses are based on Appendix 100 definitions and grading of adverse events is based on the DAIDS Table for Grading the Severity of Adult and Pediatric Adverse Events (Version 1.0, dated December 2004) grading criteria:

- As stated in Section 6.1.4 of the protocol, grading for ALT (SGPT), AST (SGOT) and total bilirubin (or direct bilirubin for women taking Atazanavir) are based on the modified grading table for hepatotoxicity in Appendix III of the protocol.
- As stated in Section 6.1.3 of the protocol, the Brief Peripheral Neuropathy Screen (BPNS) grading criteria (see Appendix II of protocol) will be used instead of the DAIDS grading criteria; signs/symptoms should be present for at least two weeks to be considered peripheral neuropathy.

3.2 Secondary Endpoints

- ***In utero* exposure and infant endpoints:**
    - Fetal death
    - Small for gestational age
    - Premature birth (< 37 weeks gestation)
    - Low birth weight (< 2500 gm)
    - Congenital anomalies
    - Grade 3 or higher clinical or laboratory AEs
        - ♦ all AEs
        - ♦ possibly, probably, or definitely related to INH/Placebo for INH
    - HIV infection
    - Hospitalization

- **Any of the following endpoints:**
    - Maternal TB (probable or confirmed – see Appendix 100 for definitions)
    - Infant TB (probable or confirmed, as defined in Appendix 100, and congenital TB, as defined using the Cantwell criteria)
    - Death of infant
    - Death of mother
    - Combined endpoint of maternal TB or maternal death
    - Combined endpoint of infant TB or infant death
    - Combined endpoint of maternal TB, maternal death, infant TB, or infant death

NOTE: TB endpoints and infant cause of death will be reviewed by a Secondary Endpoint Review Committee (SERC), which consists of maternal and pediatric TB specialists who may be

members of the study team. Refer to the Endpoint Review Process document for details of the review process. Appendix 100 will be the basis for adjudication of the TB endpoint by the SERC and is available in the CRF Appendix Codes section of the Frontier Science IMPAACT Portal; the web link to portal can be found in the MOP. Cantwell criteria for Congenital TB are in the MOP and will be collected on the P1078 CRF.

Deaths of enrolled women will receive external review by the IERC for adjudication of study treatment attribution and will also be reviewed by the SERC. The SERC will consider if the cause of death (COD) is due to unrecognized active TB (expected to be rare), from a known cause other than active TB, or from an unknown cause.

- **Occurrence of each of the following safety outcomes in HIV-infected women: during pregnancy up to: (i) delivery (antepartum) and (ii) 12 weeks postpartum**
    - Grade ≥ 3 AEs possibly, probably, or definitely related to INH/Placebo for INH or permanent discontinuation of INH/Placebo for INH due to an adverse event
    - Grade ≥ 3 AEs (all-cause)
    - Hepatotoxicity possibly, probably, or definitely related to INH/Placebo for INH
    - Hepatotoxicity (all-cause)

Note: Each of the definitions of hepatotoxicity will be studied: (i) having ≥ Grade 3 LFTs (ALT > 5 X ULN or ALT > 3 X ULN with bilirubin > 2 X ULN) or ALT > 3 X ULN with persistent, symptomatic clinical hepatitis); and (ii) using DAIDS AE grading criteria, Version 1.0.

- **Resistance outcomes:**
    - Resistance to INH in isolates of *M.tb* from mothers who develop culture-confirmed TB
    - Resistance to INH in isolates of *M.tb* from infants who develop culture-confirmed TB

3.3 Other Outcomes

- **Latent TB Diagnostic Outcome Measures**
    - IGRA assay result (positive, negative, indeterminate) in women
    - IGRA assay result (positive, negative, indeterminate) in infants
    - TST result (positive or negative) in women; TST positive if ≥ 5 mm
    - TST result (positive or negative) in infants; TST positive if ≥ 10 mm in HIV-negative infants and if ≥ 5 mm in HIV-positive infants

- **Adherence to INH/Placebo for INH among women, assessed by self-report and pill count**
    - For the self-report and pill counts, adherence will be considered as a continuous, ordinal categorical, or a binary variable. When presented as a continuous variable, adherence is defined as the percentage of expected doses/pills taken during the entire treatment period (28 weeks). When presented as an ordinal variable, excellent adherence is defined as taking at least 90% of doses/pills

during the entire treatment period; good adherence as ≥ 80% to < 90%; reasonable adherence as ≥ 60% to < 80%; and poor adherence as < 60% of doses, as reported in any of the scheduled adherence evaluations. When presented as a binary variable (completer versus non-completer), a completer is defined as having excellent adherence.

- **Population Pharmacokinetics**
  - Primary: Area under the plasma versus time curve (AUC) at each of two time points (antepartum and postpartum) in women.
  - Secondary: C_{peak}, half-life ($t_{1/2}$), volume of distribution (V_d) and apparent clearance (CL/F)

- **Pharmacogenetics**
  - Presence of genetic variants to NAT2 and 2E1, PK of INH and EFV based on sparse sampling during pregnancy and postpartum

- **Immunologic Outcome Measures**
  - Functional and phenotypic characteristics of TB specific T-cell responses in women
  - Functional characteristics of TB specific T-cell responses in infants

- **Drug Levels in Hair**
  - INH and EFV drugs levels in women's hair samples; adverse events are as defined Grade 3 or higher adverse events in women until 48 weeks postpartum

- **Risk-benefit Outcome**
  - An overall ordinal measure of clinical outcomes in the mother-infant pair, which assigns a score to the mother-infant pair according to severity of totality of clinical outcomes (clinical outcomes include maternal death, maternal TB, maternal safety events, infant death, infant TB, infant safety events); the measure will be formulated in consultation with the study team prior to release of final analysis results to the team, and details will be included in the statistical analysis plan (SAP) for this secondary analysis.

- **Maternal and infant specific TB responses and novel biomarkers**
  - Anti-mycobacterial antibody responses in maternal plasma at all weeks when samples are collected
  - Anti-mycobacterial antibody responses in infant plasma at all weeks when samples are collected
  - Inflammation and immune activation markers (and other potential biomarkers) measured from maternal plasma/PBMC samples at all available time points
  - Inflammation and immune activation markers (and other potential biomarkers) measured from infant plasma/PBMC samples at all available time points

- **Neurotoxicity**
  - Depression scores determined from the PHQ-9 instrument; responses to neurocognitive impairment questions and Pittsburgh Sleep Quality Index subset questions

4 Statistical Methods

4.1 Analysis of the Primary Objective

Safety of INH will be assessed primarily by comparing rates of Grade ≥ 3 events possibly, probably or definitely associated with INH/Placebo for INH or permanent discontinuation of INH/Placebo for INH due to an adverse reaction between women in the immediate and deferred arms by 48 weeks postpartum.

For the final primary analysis, a confidence interval for the absolute difference in primary endpoint rates will be presented as study entry gestational age-stratified Mantel-Haenszel estimates with the confidence interval computed using the method of Greenland and Robins (Biometrics 1985;41:55-68), and will use the rounded value of 5% significance level given that sequential monitoring with a Haybittle-Peto use function boundary (alpha-value) was used for interim analyses. Non-inferiority of immediate IPT in HIV-infected pregnant women with regards to safety will be demonstrated if the upper bound of this confidence interval is less than 5%. Unweighted (i.e., ignoring stratification) estimates for the difference in rates will also be reported. (Note: Although HAART use at study entry was a stratification factor in Version 1.0 of the protocol, and removed as stratification factor in Version 2.0, analyses will not stratify by HAART use at study entry because all women enrolled - except one - were on HAART at study entry.)

Data will be censored at the latest clinic visit (or contact with the participant while) on study when the endpoint status was ascertained prior to or at the earliest of the following times: (1) date when lost to follow-up for any reason, (2) date of death due to reasons not related to INH/Placebo, (3) date of permanent discontinuation of INH/Placebo as TB preventive therapy for any reason (other than an adverse reaction) including initiation of open label TB preventive therapy or treatment of presumed active TB disease, (4) date of last observed clinic visit or communication with the participant (on or before week 48 postpartum visit window). Participants who were lost to follow-up had their last date of clinic visit (or site contact) while on study before the week 48 postpartum visit window.

The primary safety analysis will be intent-to-treat, with participants analyzed according to randomized study arm.

4.1.1 Sensitivity analyses of the primary objective

The following sensitivity analyses will be performed: (i) per-protocol, which will include and compare only participants who initiated and completed the regimen to which they were randomized according to the protocol (ie. completed study treatment or discontinued due to death or toxicity); and (ii) as-treated, which will compare participants according to the actual treatment received (ie. participants in the deferred arm who started INH in the antepartum period will be analyzed in the immediate arm).

Sensitivity analyses will be done to assess the impact of coding deaths of unknown cause and all deaths as events. Supplementary analyses will include estimation of the difference in primary endpoint rates between the two arms based on simple estimates of proportions stratified by gestational age, where endpoints will be assessed based on available data (i.e. if participant discontinued the study early and did not experience an endpoint up until that point, then assume no endpoint was reached). Un-weighted (i.e., ignoring stratification by gestational age) estimates for the difference in rates will also be reported.

Multivariable analyses to adjust for maternal characteristics such as treatment arm, age, duration on HAART, CD4 count, gestational age at study entry, IGRA positivity at study entry in all women, and other potential confounders such as HBsAg serology status at entry will be performed using Poisson regression modeling.

Finally, similar analyses will be performed on all Grade ≥ 3 AEs (regardless of INH/Placebo for INH attribution) as a way to check for bias in the primary endpoint analysis.

4.2 Analysis of Secondary Objectives

Maternal TB, maternal death, infant TB and infant death will be analyzed separately and as combined events. Efficacy of INH will be assessed primarily by comparing event rates between women in the immediate and deferred arms by 48 weeks postpartum.
A 95% confidence interval for the absolute difference in event rates will be computed using study entry gestational age-stratified Mantel-Haenszel estimates. Unweighted (i.e., ignoring stratification) estimates for the difference in rates will also be reported.

Strategies for censoring will vary depending on the efficacy outcome of interest. For analyses of maternal TB, any maternal death not caused by TB will not be considered an event and any maternal death of unknown cause will be considered an event. Similarly, for analyses of infant TB, any infant death not caused by TB will not be considered an event and any infant death of unknown cause will be considered an event. Sensitivity analysis to assess the effect of this strategy includes coding death due to unknown cause as a non-event.

This will be a modified intent-to-treat analysis, with participants analyzed according to randomized study arm, excluding those who were determined to have active TB at entry.

For final analyses, supplementary analyses will include estimation of differences in these individual/combined endpoint rates between the two arms based on simple estimates for proportions.

5 Report Components

Notes:

- *Unless otherwise indicated, all outputs will be by treatment arm.*
- *Data presented on individual participants will be identified by Public ID.*
- *Validation requirements per CBAR SOPs:*
    - *All programs that generate permanent study-specific derived datasets that contribute to the generation of reports for final analyses that either contain or are supportive of the analyses of the primary or secondary objectives will be validated; all such programs will undergo code review and execution checks, and the primary safety outcome measures will have results verification (double programming)*
    - *All programs that conduct key analyses for primary or secondary objectives will be validated; all such programs will undergo code review and execution checks, and those programs that conduct the analyses for the primary objective (maternal safety) will have results verification (double programming). The following analyses are considered key: Study/Treatment Status, Study Discontinuation prior to Evaluation of Primary Outcome, Maternal Safety, Infant Safety, Maternal TB, Infant TB. The level of validation required for each of the key analyses is specified below in the respective subsections.*

5.1 Accrual

Validation: No validation of the analysis programs is required.

- Table of site activation dates and date of first enrollment
- Table of accrual by month with dates of first and last enrollment will be indicated as a footnote [STATUS]
- Table of accrual by site [STATUS]
- Table of accrual by HAART stratification factor and site [STATUS]
- Table of accrual by gestational age stratification factor and site [STATUS]
- Table of accrual by sub-cohort (intensive PK, PBMC collection, none) and site [STATUS]

5.2 Study Status

Validation: The analysis programs for Study Status and Premature Study Discontinuation will be validated, but not double-programmed.

5.2.1 Maternal

- CONSORT diagram, including number of women randomized to each treatment arm (reasons not randomized), number who initiated study treatment (reasons not initiated), and number who completed study (reasons not completed). [STATUS]
- Listing of participants with eligibility violations, including blinded ID, treatment arm, and description of violation [STATUS]
- Table of stratification errors: number (%) by strata [STATUS]
- Table of off-study reasons: number (%) [F1601]
- Table of off-study reasons: number (%) by month and phase of pregnancy (entry to < delivery, delivery to < week 12 PP, week 12 PP to week 48 PP) [F1601]
- Listing of participants not initiating randomized treatment, including blinded ID, treatment arm, and reason [ADM0021]
- Listing of participants who initiated randomized treatment more than 3 days after study entry, including blinded ID, treatment arm, and day of initiation [ADM0021]
- Duration of follow-up defined as time (weeks) from randomization date to last clinic visit (or last contact) while on study: mean (s.d.), min, max, median (Q1-Q3). [STATUS, F1601]
- If more than 10% of participants prematurely discontinue the study, generate the following:
    - Figure of time-to-last clinic visit from randomization date, with comparison by treatment group using log rank test. [STATUS, F0010]
    - Figure of time-to-last clinic visit from randomization date by gestational age stratum with comparison by treatment group using log rank test. [STATUS, F0010]

5.2.2 Infant

- Table of deliveries: number (%) by month [STATUS, EVW0292]
- Table of off-study reasons: number (%) [F1601]

5.3 Maternal Baseline Characteristics

Validation: No validation of the analysis programs is required.

5.3.1 Demographics

- Age (years): N, N missing, mean (s.d.), min, max, median (Q1-Q3); number (%) by age group (eg. <24, 24-<29, 29-<34, ≥34) with years rounded down [STATUS]
- Self-reported race/ethnicity: number (%) by category [STATUS]
- Country: number (%) [STATUS]
- Gestational age (weeks) at entry: N, N missing, mean (s.d.), min, max, median (Q1-Q3); number (%) by category (14-<24, 24-34) [OBW0006]
    - Note: Gestational age is determined based on best available method, with methods prioritized as (i) ultrasound, (ii) last menstrual period, (iii) fundal height, and may

not match the level used for stratification if the site defined best available method differently.

5.3.2 Health Status

- Entry HIV-1 RNA (copies/ml): Number (%) by category (eg. <50 cp/ml, 50-<200 cp/ml, ≥200 cp/ml) and detectable vs. undetectable (<LLQ). Note: HIV-1 RNA is performed by a "licensed assay" so the limit of quantification could differ by assay type. [RNALDMS]
- Screening CD4 counts (cells/uL): N, N missing, mean (s.d.), min, max, median (Q1-Q3); number (%) by category (eg. <350, 350-<500, 500-<650, ≥650) [LBW0054]Entry weight (kg): N, N missing, mean (s.d.), min, max, median (Q1-Q3) [F0031]
- Entry BMI (kg/m^2): N, N missing, mean (s.d.), min, max, median (Q1-Q3) [F0031,EVW0247]
    - *BMI=(weight/(height*height/10000))*
- Entry WHO clinical staging for HIV: Number (%) by clinical stage (I, II, III, IV) [PE0043]
- Prior TB treatment (>30 days): number (%) by response (yes, no, unknown) [HXW0166]
- Entry IGRA positivity: Number (%) by status (positive, negative, indeterminate) [TBW0060]
- Entry HBsAg serology: Number (%) by status (positive, negative, indeterminate) [SRW0026]
- Entry hepatitis C serology: Number (%) by status (positive, negative, indeterminate) [SRW0026]

5.3.3 Antiretroviral (ARV) Regimen

- Entry ARV regimen: number (%) by HAART category (HAART, not on HAART) and by actual regimen [PE0412]
- Efavirenz-containing ARV regimen at entry: number (%) by EFV use (yes, no) [PE0412]

5.3.4 Laboratory Values

- ALT: number (%) by grade; mean (s.d.), min, max, median (Q1-Q3)
- AST: number (%) by grade [PE6817]
- Total Bilirubin: number (%) by grade [PE6817]
- ANC (cells/mm^3): number (%) by grade [PE6812]
- Hemoglobin (g/dl): number (%) by grade [PE6812]
- Platelet count (cells/mm^3): number (%) by grade [PE6812]

Note: Grading will be based on Appendix III of protocol for ALT, AST and total bilirubin and by DAIDS Version 1.0 toxicity grade otherwise. Baseline laboratory evaluations were done at screening, except for ALT, which was done at screening and entry. Baseline ALT is calculated as the mean of screening and entry values if both are available.

5.4 Maternal Safety

Validation: The analysis programs in this section will be validated, but not double-programmed.

- Tables of new grade 3 or 4 labs and signs/symptoms: number (%) by SOC and grade. [STATUS, PE6812, PE6817, PE6832]
- Table of new grade 3 or 4 diagnoses: number (%) by MedDRA code [EVENTS].
- Table of new grade 3 or 4 primary adverse events (labs, signs/symptoms, and diagnoses): number (%) by MedDRA code [EVENTS]
- Table of serious adverse events (SAEs): number (%) by MedDRA code [EVENTS]

Note: The highest grade experienced by a given participant for each safety event is included.

5.5 Maternal Hepatotoxicity

Validation: The analysis programs in this section will be validated, but not double-programmed.

- Table of grade 1, 2, 3, and 4 LFTs (ALT, AST, TBIL): number (%) by highest grade [PE6817]
- Table of hepatotoxicity by type, assuming protocol grading and DAIDS grading, separately: number (%) with Fisher's exact tests with mid-*p* adjustment (Lydersen et al, 2009, Statistics in Medicine) [PE6817]:
    - Grade 3 ALT (highest grade)
    - Grade 4 ALT (highest grade)
    - Grade 3 ALT, AST, or TBIL (highest grade)
    - Grade 4 ALT, AST, or TBIL (highest grade)
    - Grade ≥2 ALT and grade ≥2 TBIL
    - Grade ≥2 ALT with symptomatic clinical hepatitis
    - Any hepatotoxicity (defined as any of the above)

    Note: Symptomatic clinical hepatitis will be identified based on blinded review (pooled over study arms) by selected protocol team members (at end of study follow-up and prior to data freezing for final analysis). Substitute direct bilirubin for TBIL for participants on atazanavir.

- Create the following by treatment arm and efavirenz use at entry, separately [PE6817, ADM0040]:
    - Peak ALT: min, max, mean (SD), median (Q1-Q3), Wilcoxon rank sum test
    - Table of participant ALT levels (U/L) at scheduled study visits by stage of pregnancy (weeks AP/PP): N, mean (s.d.), median (Q1-Q3)
    - Table with number (%) of participants by ALT grade at scheduled study visits and stage of pregnancy (weeks AP/PP)

    - Plot of median (Q1,Q3) ALT level at scheduled study visits by stage of pregnancy (weeks AP/PP)
- Table of estimates (with 95% confidence intervals) of difference in incidence rates between arms, based on stratified Mantel-Haenszel estimates (primary) and simple proportions (sensitivity) at week 48 postpartum, for the following endpoints [STATUS,PE4005,TRAC]:
    - Any hepatotoxicity, possibly, probably, or definitely related to study treatment
    - Any hepatotoxicity, all cause
- Kaplan-Meier plot of time to endpoint from randomization date by study arm and by gestational age, with number of events and number at risk for the following endpoints
    - Any hepatotoxicity, possibly, probably, or definitely related to study treatment
    - Any hepatotoxicity, all cause
- Table of analysis of difference in endpoint rates between arms up to (i) delivery and (ii) 12 weeks postpartum, based on stratified Mantel-Haenszel estimates (primary) and simple proportions (sensitivity), for the following endpoints [STATUS,PE4005,TRAC]:
    - Any hepatotoxicity, possibly, probably, or definitely related to study treatment
    - Any hepatotoxicity, all cause

5.6 Peripheral Neuropathy

Validation: The analysis programs in this section will be validated, but not double-programmed.

- Table with ≥grade 1 peripheral neuropathy: number (%) by highest grade [NE6056]
- Table of BPNS final score: min, max, mean (SD), median (Q1-Q3), Wilcoxon rank sum test [NE6056]

5.7 Maternal Mortality

Validation: The analysis programs in this section will be validated, but not double-programmed.

- Listing of maternal deaths including public ID, treatment arm, weeks on study, stage of pregnancy, site assessment of primary COD and treatment attribution, and IERC treatment attribution [STATUS, PE1414, TRAC]
- Listing of maternal deaths including the following peak lab values: ALT, AST, total bilirubin, alkaline phosphatase, GGT, INR, total protein, and albumin [PE6812, PE6817]

5.8 Treatment Intolerability

Validation: The analysis programs in this section will be validated, but not double-programmed.

- Table of treatment discontinuation: number (%) of enrolled women discontinuing study treatment prematurely (discontinued study treatment prior to week 40 postpartum, which is the end of study treatment period for both arms) by reason [PE4005]

- Table of permanent treatment discontinuation due to toxicity: number (%) of enrolled women discontinuing study treatment prematurely due to toxicity by type of toxicity [PE4005, TXW0221]
- Listing of participants who permanently discontinued study drug due to toxicity but did not have a grade 3 or higher AE related to study treatment, including: blinded ID; treatment arm, actual week of treatment discontinuation, and off-treatment reason [PE4005, TRAC]

5.9 Primary Endpoints and Analysis

Validation: The analysis programs in this section will be validated, and primary analysis will be double-programmed.

- Cross-tabulation of participants with at least one adverse event at least possibly related to study treatment based on IERC assessment and participants with at least one adverse event at least possibly related to study treatment based on site assessment: number (%) [TRAC]
- Table of most severe safety outcome related to study treatment among participants experiencing primary safety endpoints: number (%) of each of the following [STATUS,PE4005,TRAC]:
    - Grade 3 AE related to study treatment
    - Grade 4 AE related to study treatment
    - Death related to study treatment
    - Permanent discontinuation of study treatment due to toxicity but no grade 3 or higher AE related to study treatment
- Table of primary analysis of difference in endpoint rates between arms, based on stratified Mantel-Haenszel estimates at week 48 postpartum (primary) and simple proportions (sensitivity), for the following endpoints [STATUS,PE4005,TRAC]:
    - Primary safety endpoint
    - Any Grade ≥ 3 AE (all cause)
- Table of per-protocol and as-treated sensitivity analyses of difference in endpoint rates between arms, based on stratified Mantel-Haenszel estimates at week 48 postpartum and simple proportions, for the following endpoints [STATUS,PE4005,TRAC]:
    - Primary safety endpoint
    - Any Grade ≥ 3 AE (all cause)
- Table of sensitivity analysis coding deaths of unknown cause and all deaths as primary endpoints (if death or a prior AE experienced by the participant who died were not already considered to be a primary endpoint)
- Kaplan-Meier plot of time to endpoint from randomization date by study arm and by gestational age, with number of events and number at risk for the following endpoints
    - Primary safety endpoint
    - Any Grade ≥ 3 AE (all cause)
- Table of analysis of difference in endpoint rates between arms up to (i) delivery and (ii) 12 weeks postpartum, based on stratified Mantel-Haenszel estimates (primary) and simple proportions (sensitivity), for the following endpoints [STATUS,PE4005,TRAC]:

  - Primary safety endpoint
  - Any Grade ≥ 3 AE (all cause)
- Multivariable Poisson regression model for the primary safety endpoint adjusted for treatment arm and the following baseline characteristics: maternal age, duration on HAART (from start date of HAART to randomization date), CD4 count, gestational age, IGRA positivity, and HBsAg serology status

Per NIH policy for Phase III and pivotal Phase II and IV studies, NIH requires primary analyses of treatment comparisons to be summarized by sex and by race and treatment interactions with sex and race to be tested. These analyses are required so do not represent multiple comparisons and are presented in the primary study analysis regardless of power issues. As P1078 maternal analyses are restricted to women, the primary analysis will be presented by race but not by sex.

- Table of primary analysis of difference in primary safety endpoint rates between arms, based on stratified Mantel-Haenszel estimates at week 48 postpartum, by race [STATUS,PE4005,TRAC]
- Logistic regression of the primary safety endpoint on the interaction term for treatment and race

5.10 *In Utero* Exposure and Infant Outcomes

Validation: The analysis programs in this section will be validated, but not double-programmed.

Note: Exact categorical tests (with mid-p adjustment for the case of Fisher's exact tests) will be employed to compare the difference in outcomes between arms.

5.10.1 Summary of pregnancy outcomes

- Table with number (%) of pregnancies by outcome type (eg. live birth, stillbirth, spontaneous abortion, induced abortion, discordant birth outcomes) and number of fetuses (singleton, twins)

5.10.2 Infant baseline characteristics

- Gender: number (%) by category [EVW0292]
- Gestational age at birth (weeks): N, N missing, min, max, median (Q1-Q3); number (%) by category (<34, 34 -<37, ≥37)
  - Note: The calculation of gestational age at birth will be based on three sources, prioritized as follows: newborn exam [PE5896], pregnancy outcome evaluation [EVW0292] and mother's gestational age at enrollment [OBW0006] plus number of weeks to delivery.
- Birth weight (grams): N, N missing, min, max, median (Q1-Q3); number (%) by category (<1500, 1500 -<2500, ≥2500) [PE5896]

- Apgar score at 1 minute: N, N missing, min, max, median (Q1-Q3); number (%) by category (0-3, 4-6, 7-10) [PE5896]

5.10.3 Summary of adverse pregnancy outcomes

Note: The evaluations of small for gestational age, premature birth, low birth weight and congenital anomalies are based on live births. Multiple births with both live birth and stillbirth outcomes will be counted as pregnancies with stillbirths and other adverse pregnancy outcomes that may have been observed. Pregnancy outcome of induced abortion will be excluded from adverse pregnancy outcome summaries.

- Table with number (%) of pregnancies resulting in each of the following adverse pregnancy outcomes:
    - Fetal demise (stillbirth, spontaneous abortion)
    - Premature birth (<37 weeks gestation)
    - Low birth weight (<2,500 grams)
    - Congenital anomalies
    - Any adverse pregnancy outcome excluding congenital anomalies (defined as any of the above excluding congenital anomalies)
    - Any adverse pregnancy outcome (defined as any of the above)

- Table with number (%) of pregnancies resulting in each of the following severe adverse pregnancy outcomes:
    - Fetal demise (stillbirth, spontaneous abortion)
    - Premature birth (<34 weeks gestation)
    - Low birth weight (<1,500 grams)
    - Any severe adverse pregnancy outcome (defined as any of the above)

5.10.4 Summary of HIV infection status

- Table with infant HIV infection status (positive, negative, indeterminate, unknown)

  Note: To be reviewed at end of study follow-up and prior to data freezing for final analysis by selected protocol team members (pooled over study arms).

5.10.1 Summary of safety and mortality

- Tables of new grade 3 and 4 labs and signs/symptoms: number (%) by SOC and grade [STATUS, PE6812, PE6817, PE6832]
- Table of new grade 3 and 4 diagnoses: number (%) by MedDRA code [EVENTS].
- Table of new grade 3 and 4 clinical and laboratory events: number (%) by MedDRA code [EVENTS]
-
- Listing of infant/neonatal deaths, including blinded ID, treatment arm, age at death (weeks on study), gestational age at birth, primary cause of death, narrative of etiology of death [PE1414]

- Table of infant safety summary: number (%) with any grade 3 or 4 adverse event; grade 3 or 4 adverse event definitely, probably or possibly related to study treatment based on site assessment; hospitalization; infant/neonatal death (overall, within 0-7 days of birth, within 0-28 days of birth) [EVENTS,PE9905,PE1414]

5.11 Maternal and Infant TB

Validation: The analysis programs in this section will be validated, but not double-programmed.

Note: Mother-infant pairs will be excluded if the mother had TB diagnosis at study entry.

- Listing of active TB events, including blinded ID, treatment arm, event week, site diagnosis, SERC decision, event grade and INH resistance testing results (if available), shown separately for mothers and infants
- Table of SERC-assessed TB and death endpoints: number (%) by arm of the following:
    - Maternal TB cases by type (PULMONARY: 220021-CONFIRMED, 220022-PROBABLE; EXTRAPULMONARY: 220031-CONFIRMED, 220032-PROBABLE)
    - Infant TB cases by type (PULMONARY: 220021-CONFIRMED, 220022-PROBABLE, POSSIBLE; EXTRAPULMONARY: 220031-CONFIRMED, 220032-PROBABLE; CONGENITAL TB.
    - Death of infant
    - Death of mother
    - Maternal TB or maternal death
    - Infant TB or infant death
    - Maternal TB, maternal death, infant TB, or infant death

As event numbers warrant:

- Table of intent-to-treat analysis of difference in endpoint rates between arms, based on stratified Mantel-Haenszel estimates at week 48 postpartum (primary) and simple proportions (sensitivity), stratified by gestational age, for the following endpoints:
    - Maternal TB
    - Infant TB
    - Death of infant
    - Death of mother
    - Maternal TB or maternal death
    - Infant TB or infant death
    - Maternal TB, maternal death, infant TB, or infant death

- Kaplan-Meier plot of time to endpoint from randomization date by study arm and by gestational age, with number of events and number at risk for the following endpoints:
    - Maternal TB
    - Infant TB
    - Death of infant

- Death of mother
- Maternal TB or maternal death
- Infant TB or infant death
- Maternal TB, maternal death, infant TB, or infant death

5.12 Changes in ARV regimen

Validation: The analysis programs in this section will be validated, but not double-programmed.

- Table of ARV discontinuation: number (%) discontinuing at least one component of baseline ARV regimen by category (EFV-based regimen vs. non-EFV ARV regimen) [PE0421]
- Table of changes in EFV use: number (%) with the following changes in EFV use, by phase of pregnancy (entry to < delivery, delivery to < week 12 PP, week 12 PP to week 48 PP) [STATUS, PE0421]:
  - Switched from an EFV-containing regimen at entry to a non EFV-containing regimen
  - Switched from a non EFV-containing regimen at entry to an EFV-containing regimen

5.13 Vital Status of Participants Lost to Follow-up

Validation: The analysis programs in this section will be validated, but not double-programmed.

- Table of number (%) of participants lost to follow-up with vital status data and source of information by category (participant, family member, neighbor, medical records, other) [VSW0024]
- Table of number (%) of mothers who experienced the following since last seen on study [VSW0024]:
  - Death
  - Hospitalization
  - TB diagnosis
  - Delivery by pregnancy outcome
- Table of number (%) of infants who experienced the following since last seen on study [VSW0024]:
  - Death
  - Hospitalization
  - TB diagnosis

5.14 Report Appendices

- Listing by arm of enrolled women who experienced primary safety endpoint, including: blinded ID; adverse event and grade associated with safety endpoint; week of event;

whether treatment was temporarily discontinued, permanently discontinued, or not; week of treatment discontinuation [STATUS, PE4005, TRAC]