This is ACTG A5263/AMC 066 SAP Version 2.0 with names of authors, names of publication writing team members and analysis timeline redacted.

ACTG A5263/AMC 066

A Randomized Comparison of Three Regimens of Chemotherapy with Compatible Antiretroviral Therapy for Treatment of AIDS-KS in Resource Limited Settings

ClinicalTrials.Gov ID: NCT01435018

Statistical Analysis Plan

Version 2.0

September 26, 2018

Table of Contents

1 Introduction

The Primary Statistical Analysis Plan (SAP) describes the proposed content for the primary statistical analysis report of ACTG A5263, which addresses the primary and secondary objectives of the study. This document also describes the primary and secondary outcome measures for which results will be posted on ClinicalTrials.gov. The Primary SAP outlines the general statistical approaches that will be used in the analysis of the study and has been developed to facilitate discussion of the statistical analysis components amongst the study team, and to provide agreement between the study team and statisticians regarding the statistical analyses to be performed and presented in the primary analysis report.

Because the study was closed prematurely (triggered by a DSMB review and recommendation), analyses for the primary analysis report will be initiated as soon as possible; no analysis timeline was generated. In addition, all primary and secondary outcomes outlined in this SAP will be submitted to ClinicalTrials.gov within 1 year of the primary completion date (PCD). The PCD has been adjusted to be the date of the last interim review, which occurred on March 13, 2018; only data as of this date will be included in the primary analysis.

Additional specifications and details on other key elements of the primary analysis report are provided in a separate document, the Analysis Implementation Plan (AIP).

1.1 Key Updates to the SAP

Version 1.0 of the SAP (finalized on February 16, 2016) included details on interim analysis methods and planned analyses for the primary outcome. The SAP was updated after the final interim review, but prior to the final analysis, and includes the following changes:

- Reformat to be compliant with new SOP requirements
- Add details on study history related to interim analysis methods
- Move specifics on basic report elements to the AIP
- Modify approach for primary analysis to address statistical issues related to premature study closure
- Add sections on secondary outcomes (not previously discussed in prior SAP).

2 Study Overview

Version 1.0 of the study protocol was finalized on October 14, 2010. Due to a world-wide shortage of the active-control drug, doxorubicin HCL liposome, the protocol was amended to include a different active-control, paclitaxel. Participants began enrollment under version 2.0 of the study protocol (finalized on August 7, 2012).

2.1 Design

A5263/AMC066 is a randomized, prospective, active-controlled, clinical trial designed to compare three regimens of chemotherapy with compatible antiretroviral therapy for the treatment of advanced AIDS-KS in resource-limited settings. Randomization (in each study step) is stratified by CD4+ cell count (<100 or ≥ 100 cells/mm^3) and country. The trial was designed to evaluate whether there is sufficient evidence to conclude if bleomycin and vincristine (BV) plus ART is noninferior to paclitaxel (PTX) plus ART, or if oral etoposide (ET) plus ART is non inferior to PTX plus ART.

The study was designed as a four-step study with participants randomized in the following manner:

Step 1: Participants were randomized 1:1:1 to ET+ART, BV+ART, or PTX+ART

Step 2: Repeat of treatment received in Step 1 if conditions outlined in protocol were met

Step 3: Participants were randomized 1:1 to the two remaining treatment arms not utilized in Step 1 (or Step 1 and 2) if certain conditions outlined in the protocol were met

Step 4: Participants were assigned to the remaining treatment not utilized in prior steps if conditions outlined in the protocol were met

Under protocol version 2.0, the study planned to enroll a total of 706 HIV-infected men and women at least 18 years of age with the following characteristics:

- Biopsy-confirmed diagnosis of KS
- Naïve to ART (allowed within 42-day window prior to study entry), chemotherapy, and radiation therapy
- KS stage T1
- Persons with chronic, acute, or recurrent infections that are serious, in the opinion of the site investigator, must have completed at least 14 days of therapy prior to study entry and must be clinically stable

All participants were to be followed for 5 years after randomization or assignment to the last step of entry, with total follow up ranging from 5 to 7 years based on planned enrollment.

The protocol was substantially modified after the March 2016 interim review. At this review the DSMB recommended that enrollment to the ET+ART arm be stopped on all steps. Protocol version 3.0 was finalized on December 7, 2016, and incorporated all prior amendments, edits necessary to implement the closure of the ET+ART arm, and a modified sample size. The modified sample size was based not only on the closure of the ET+ART arm, but also on the ACTG SASC recommendations to reduce study power from 88% to 80% for feasibility reasons (i.e., to reduce sample size and decrease the time needed to complete accrual).

Under protocol version 3.0, the study planned to enroll a total of 446 participants (386 randomized to BV+ART or PTX+ART, 60 participants previously randomized to ET+ART).

Duration of follow-up was also modified with version 3.0 based on recommendations from the ACTG SASC. With this change, all participants who ever received ET on study would be followed for 144 weeks after the start of their last cycle of ET; all other participants would be followed for 96 weeks after randomization or assignment to their last study step of entry.

Study visit schedules follow the same format on each step and remained unchanged with the protocol modifications. Visits occurred every 3 weeks through week 48, and then occurred every 12 weeks until week 96. Study week windows occurred ± 7 days for all visits through week 48, and then ± 14 days for visits through week 96. For those who received ET, all additional visits through end of follow-up occurred ± 12 weeks.

2.2 Power and Sample Size

The study planned to enroll 706 participants under version 2.0 of the protocol. With this sample size there was 88% power to demonstrate non-inferiority assuming the progression-free survival (PFS) rate in each arm was 65%; each experimental arm (ET+ART, BV+ART) is to be contrasted separately with PTX+ART with Type-I error rate control at the contrast level (trial-wise Type I error will be larger than 5%).

This sample size was determined based on the following assumptions:

- 15% non-inferiority margin
- Two interim analyses and one final analysis would be conducted
    - Interim analyses would occur after 33% and 67% information
    - Lan-DeMets spending function was used corresponding to the O'Brien-Fleming boundary
- One-sided significance level α=0.025
- 1:1:1 assignment fraction
- Inflation rate of 10% for loss-to-follow-up

The sample size was modified in version 3.0 of the protocol and reflects the changes associated with the closure of the ET+ART arm and reduced power. With the same above-stated assumptions the targeted sample size for the BV+ART and PTX+ART arms was 386 total participants. With 60 participants already enrolled to the ET+ART arm, the total planned sample size for the study under version 3.0 was 446.

2.3 Hypotheses

1. ET plus ART is non-inferior to PTX plus ART for initial treatment of advanced stage AIDS-related Kaposi's Sarcoma (AIDS-KS).

2. BV plus ART is non-inferior to PTX plus ART for initial treatment of advanced stage AIDS-KS.

2.4 Objectives

The Primary SAP addresses the following primary and secondary objectives listed in the study protocol. Secondary objectives addressing quality of life [Protocol Objective 1.2.2] will be analyzed outside of SDAC and those analyses will be outlined in a separate SAP. The remaining study objectives in the protocol will be addressed in subsequent analysis plans [Protocol Objectives: 1.2.8, 1.2.9, and 1.2.11]; these are related to long-term safety follow-up (which is ongoing) and laboratory-based objectives that are pending sample mobilization and testing. No statistical analysis will be conducted to address Protocol Objective 1.2.10, which was to encourage donations of excess biopsy materials.

2.4.1 Primary Objective

1. To compare the clinical efficacy of two regimens, etoposide (ET) plus ART and bleomycin plus vincristine (BV) plus ART, to paclitaxel (PTX) plus ART for initial treatment of advanced stage AIDS-related Kaposi's sarcoma (AIDS-KS). [Protocol Objective 1.1]

2.4.2 Secondary Objectives

1. Compare Kaposi's sarcoma (KS) tumor response in persons randomized to ET plus ART, BV plus ART, and PTX plus ART. [Protocol Objective 1.2.1]

2. Compare the safety and toxicity in persons randomized to ET plus ART, BV plus ART, and PTX plus ART. [Protocol Objective 1.2.3]

3. Compare suppression of plasma HIV-1 RNA and changes in CD4+ lymphocyte cell count in persons randomized to ET plus ART, BV plus ART, and PTX plus ART. [Protocol Objective 1.2.4]

4. Compare adherence to ART in persons randomized to ET plus ART, BV plus ART, and PTX plus ART. [Protocol Objective 1.2.5]

5. Compare incidence of peripheral neuropathy (PN) and symptomatic peripheral neuropathy (SPN) in persons randomized to ET plus ART, BV plus ART, and PTX plus ART. [Protocol Objective 1.2.6]

6. Evaluate clinical efficacy of the remaining regiment after failure of the initial regimen. [Protocol Objective 1.2.7]

2.5 Outcome Measures

2.5.1 Primary Outcome Measure

1. Progression free survival (PFS) defined as a lack of IERC-confirmed KS progression, death, entry to an additional step, or loss to follow-up prior to week 48.

   All potential KS disease progressions were reviewed by an independent endpoint review committee (IERC), which consisted of experienced investigators who were not members of the protocol team and were blinded to treatment assignment. Sites submitted a request for the IERC to review a potential KS progression by entering CRF EVW0232 (KS Evaluation) and EVW0233 (KS Follow-Up Evaluation) and using the IERC Real-Time Clinical Review interface to submit clinical findings and photographic record; the IERC responded within 48 hours.

2.5.2 Secondary Outcome Measures

1. Death by week 48
2. IERC-confirmed KS progression by week 48
3. AIDS-defining event by week 48
4. HIV-1 RNA virologic failure by week 48

   Virologic failure (VF) is defined as two-successive measurements of plasma HIV-1 RNA ≥ 1000 copies/mL at week 12 to 24 or, HIV-1 RNA ≥ 400 copies/mL at week 24 or later. The date of VF will correspond to the timing of the first of the two measures; if there is not a confirmatory viral load result then the conservative approach of classifying the single measure has a VF will be used

5. KS-IRIS by week 48

   KS-IRIS is defined as any IERC-confirmed KS progression that occurs within 12 weeks of initiation of ART that is associated with an increase in CD4+ cells of at least 50 cells/mm^3 above the study entry value and/or decrease in the HIV-RNA level by at least 0.5 log below the study entry value prior to or at the time of documented IERC-confirmed KS progression.

6. Objective Response (OR) during Step 1

   Objective Response is defined as either a complete response (CR) or partial response (PR).

7. Duration of Objective Response (OR) during Step 1

   Duration of OR is defined as time from initial Objective Response to time of IERC-confirmed KS progression among those with an OR

8. IERC-confirmed KS progression, death, or AIDS-defining event by week 48
9. IERC-confirmed KS progression, death, AIDS-defining event, or VF by week 48
10. IERC-confirmed KS progression, death, AIDS-defining event, VF, or KS-IRIS by week 48
11. Time to IERC-confirmed KS-progression or death, whichever is earlier
12. Time to death

13. Change in KS treatment by week 48

    Change in KS treatment is defined as stopping randomized chemotherapy and initiating a different chemotherapy, regardless of the reason.

14. Occurrence of post-entry Grade 3 or higher signs and symptoms and laboratory toxicities

15. Occurrence of treatment-related toxicities and adverse events (AEs)

16. Occurrence of peripheral neuropathy (PN)

17. Occurrence of symptomatic PN (SPN)

18. Change in CD4+ lymphocyte cell count from entry to weeks 12, 24, and 48

19. Adherence to ART

    ART adherence is based on participant recall and is defined as not missing ART for the month prior to the current visit

20. Presence of oral KS at entry

21. IERC-confirmed KS progression, dose-limiting toxicity, VF, AIDS-defining events, objective response, or death during Step 2

22. IERC-confirmed KS progression, dose-limiting toxicity, VF, AIDS-defining events, objective response, or death during Step 3

23. IERC-confirmed KS progression, dose-limiting toxicity, VF, AIDS-defining events, objective response, or death during Step 4

2.6 Interim Monitoring

The NIAID/DAIDS Co-infections and Complications Data and Safety Monitoring Board (CC-DSMB) was responsible for the interim reviews of A5263/AMC066, which occurred either annually or biannually and consisted of administrative, safety, and efficacy summaries. Summary reports were prepared for each review and included details on study conduct, baseline characteristics, safety and tolerability, and endpoint data by randomized treatment arm. Prior to the start of enrollment, the CC-DSMB requested that a summary of primary endpoint data by randomized treatment arm be provided in all DSMB reports. Although not pre-stated in the protocol, Kaplan-Meier methods with Greenwood's formula for the variance were used to obtain PFS rates by calculating the survival probabilities for each treatment arm at week 48. This approach was used to account for differential follow-up time by censoring observations at the most recent reported dates of contact if they had not experienced an event.

The study planned to have two interim analyses and one final analysis, the timing of which were determined using a Lan-DeMets spending function corresponding to the O'Brien-Fleming boundary. The final analysis would consist of constructing a 95.37% CI for primary endpoint (reported as a 95% to account for 95% simultaneous coverage probability given the interim analyses) and interim analyses were planned after 33.3% and 66.7% of the planned enrollment had reached week 48 (33% and 67% information), which corresponded to constructing 99.98% and 98.80% confidence intervals for the primary endpoint.

At the second interim review (September 2015), with consultation with ACTG SDAC leadership, the decision was made to provide formal treatment arm comparisons at an earlier-than-planned time point (i.e. before 33% information) because of observed differences between treatment arms; 99.98% CIs were used for this review. Henceforth, the CC-DSMB requested all future reviews contain formal between-arm comparisons, and confidence intervals were planned to be based on the available statistical information at the given review. This change in strategy for interim reviews was incorporated into the statistical section of protocol version 3.0. Specifically, it noted that formal between-arm comparison would occur at each interim CC-DSMB review with corresponding CIs being based on the estimated percent information available at that review. And, that the primary analysis would consist of constructing a CI for the primary endpoint that is adjusted using the Lan-DeMets approach to account for the multiple interim analyses in order to provide 95% simultaneous coverage over the multiple analyses.

However, because there was still limited % information at the 2016 reviews (see Table below) and O'Brien-Fleming stopping guidelines require a very extreme p-value (and CI coverage) before stopping is considered, 99.9% confidence intervals were used at these reviews. This approach is analogous to capping p-values at 0.001 and is similar to the use of Haybittle-Peto guidance for analyses conducted at an early information time. All subsequent reviews based the confidence intervals on % information.

A summary of the alpha-spending from each DSMB reviews is provided for reference.

Review	Estimated % Info	CI	Alpha Spent	Total Alpha Spent	
				ET v PTX	BV v PTX
Sept 2015	7%	99.98%	0.0002	0.0012	0.0104
March 2016	13%	99.9%	0.001		
Sept 2016	22%	99.9%	0.001		
Sept 2017	35%	99.98%	0.0002		
March 2018	60%	99.2%	0.008		

3 General Analysis Considerations

Study (Step 1) entry, denoted as *Week 0*, is defined as the date of the entry visit recorded on Study Initiation Case Report Form (CRF). Entry to other study steps is defined as the date of randomization to the respective step and is denoted as *Week R*.

For summaries that occur at specific study weeks (e.g. CD4 cell count, HIV-1 RNA levels), the following week windows are used:

Week 0:	Days [0, 1]
Week 12:	Weeks [10, 14]
Week 24:	Weeks [22, 26]
Week 36:	Weeks [34, 38]
Week 48:	Weeks [46, 50]
Week 60:	Weeks [58, 62]
Week 72:	Weeks [70, 74]
Week 84:	Weeks [82, 86]
Week 96:	Weeks [94, 98]

Loss-to-follow-up, as stated in the primary outcome, is defined as being off-study for an unknown reason. This was not explicitly defined in the protocol, but was decided prior to the first interim review with consultation with the team. As stated in the protocol, missing data (i.e., loss to follow up) was included in the primary outcome because of the advanced disease state of most enrolled participants and most missing data from international sites is due to participant death (ACTG meeting 2008 presentation).

All data listings will use CBAR *PUBLICID* and will have calendar date information redacted and will instead be summarized in terms of study week.

Descriptive summaries of continuous baseline characteristics will include number of observations, number of observations missing, median, IQR, and min/max; descriptive summaries of other continuous variables will use mean and standard deviation, or median and IQR depending upon the corresponding analysis method. The frequency and percentage will be used for all categorical variables.

Two different freeze dates will be used to construct the final analysis database. The first date, March 10th 2016, corresponds to the DSMB review date at which the decision was made to close enrollment to the ET+ART arm; this will be the freeze date for the ET+ART data. The second date, March 13th 2018, corresponds to the DSMB review date at which the decision was made to close enrollment to the remaining two arms (BV+ART and PTX+ART); this will be the freeze date for the BV+ART and PTX+ART data. All descriptive by arm summaries will include all available data as of the relevant freeze dates; the ET+ART data will be from a different freeze date than the other two arms. Formal treatment group comparisons are slightly more nuanced. Comparisons of BV+ART versus PTX+ART will use all data as of the March 13th 2018 freeze date. Formal treatment group comparison of ET+ART versus PTX+ART will use all data as of the March 10th 2016 freeze date; this will require that the PTX+ART data be further restricted to only include data from the same time period as the ET+ART arm.

The analysis population for the final analysis will include all randomized participants who were eligible for the study and who started their randomized chemotherapy; only a handful of participants were ineligible or did not initiate their assigned chemotherapy. All analyses will be conducted among this population; treatment arm comparisons will contrast initially randomized treatment arms, regardless if chemotherapy was stopped or modified. A second censored analysis will also be conducted, which censors participants if they initiate different study chemotherapy; this analysis restricts follow-up to time on the initially randomized treatment arm. For the censored analysis, participants will be censored at the time they start a different study chemotherapy (Step 3 randomization), as appropriate for the specific outcome measure. Censoring will not occur if treatment is stopped prematurely without starting a new chemotherapy.

All primary and secondary analyses outlined in this SAP will use a 5% alpha. Each experimental arm (ET+ART and BV+ART) will be contrasted separately with PTX+ART, with each contrast using a 5% alpha. This decision, to control the Type-I error rate at the contrast level (instead of at the trial level) was based on team input at the time of protocol design. Additional details are discussed in the statistical methods section.

4 Statistical Methods

4.1 Analyses of the Primary Outcome

4.1.1 Analysis Considerations

As noted in previous sections, the original plan for interim monitoring was modified during the course of the study, with formal treatment arm comparisons occurring earlier than planned. In addition, the ET+ART arm was prematurely closed at an early interim review and the overall study was terminated prior to the study fully enrolling. Both study modifications were due to inferiority of the experimental arms relative to PTX+ART. Taking into consideration this and the fact that only a small amount of alpha was spent at the interim reviews, the decision was made to use a 5% alpha for the primary analysis implemented by 95% nominal confidence intervals.

Along with the nominal 95% CIs calculated for the analyses outlined in this SAP, the results (CIs) included in the reports that led the DSMB to recommend action (i.e., closing the ET+ART arm and closing the study) will also be provided in the analysis report. These results will be taken directly from the prior interim reports; no additional analysis will be conducted to generate these intervals.

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 and do not represent multiple comparisons and will be presented in the primary study analysis regardless of power issues. The majority of participants enrolled into the study are black and therefore only analyses by sex and treatment interactions with sex will be examined.

4.1.2 Primary Outcome Methods

Within arm estimates of the PFS rate at week 48 will be obtained using Kaplan-Meier methods with Greenwood's formula for the variance (see below), and are determined by calculating the survival probabilities at week 48. Events are defined as the first occurrence of any component of the primary endpoint within the first 48 weeks. Primary endpoint components include IERC-confirmed KS progression, death, entry to another step, or loss-to-follow-up for an unknown reason. Participants could experience more than one component of the primary endpoint, but only the timing of the first component determines if they are classified as an event. Participants who have not experienced an event in the first 48 weeks will be censored at their date of last reported contact (if their total follow up time is less than 48 weeks) or at week 48 (for those with more than 48 weeks of follow up). Time is calculated as the difference from week 0 (step 1 entry) to either the event or censor date.
Summary tables describing the frequency and percentage of each component will be provided overall and by treatment arm, as well as a summary of endpoint status (i.e. frequency and percent of participants with an event, without an event and follow-up < 48 weeks, and without an event and follow-up > 48 weeks). Additional details will be provided in the AIP.

The absolute difference in 48-week PFS rates between ET+ART and PTX+ART, and BV+ART and PTX+ART will be calculated separately. The 48-week survival probabilities and two-sided 95% CIs for each difference will be determined overall (not adjusting for stratification factors) and with adjustment (weighted) for randomized stratification factors; the inverse of the stratum-specific variance will be used for the stratum weights. Adjustments will be done separately for each stratification factor (CD4+ cell count and country); an analysis that simultaneously adjusts for both stratification factors will be examined, however, this may not be possible due to the limited number of participants in some strata. The differences and corresponding CIs will be plotted in a forest plot with a reference line for the NI margin (15%).

Subgroup analyses will examine the difference in PFS rates within post-hoc identified subgroups. The subgroups considered include those with visceral disease versus those without visceral disease at entry, and those with screening CD4 cell counts <100 versus ≥ 100 cells/mm^3 (stratification factor).

Two sensitivity analyses will also be conducted. The first sensitivity analysis will analyze the primary outcome among all randomized participants regardless if they started their randomized chemotherapy or were ineligible for the study. The second sensitivity analysis will consider an expanded definition of loss-to-follow-up for the primary endpoint. This modified definition will include all non-administrative off-study reasons (i.e., will not include site or study closure as loss-to-follow-up).

Because participants are classified as an event if they enter another study step, the censored analysis will not be done for the primary analysis.

4.2 Analysis of Secondary Outcomes

4.2.1 Analysis Considerations

All analyses of secondary outcomes will use a Type-I error rate of 5%; no secondary outcomes were examined during interim monitoring so prior alpha-spending considerations are not necessary for secondary outcomes. As previously mentioned in the general analysis considerations section, all analysis will be among the defined analysis population. A censored analysis will also be done for most secondary outcomes, which censors participants at the time they start different study chemotherapy (Step 3 randomization date or other); details on exceptions are denoted in this section. For those who started different study chemotherapy due to entering Step 3, the Step 3 randomization date will be used for the censor date, for those who initiated a different study chemotherapy while on Step 1, the start date of the chemotherapy will be used as the censor date (as recorded on the concomitant medications case report form).

Prior to analysis, the team identified a list of conditions that are classified as AIDS-defining events from a list of diagnoses from Appendix 60, the Diagnoses Appendix used for this study. The list of diagnosis codes and corresponding conditions are located in the AIP.

Virologic failure (VF) is defined as two-successive measurements of plasma HIV-1 RNA ≥ 1000 copies/mL at week 12 to 24 or, HIV-1 RNA ≥ 400 copies/mL at week 24 or later. The date of VF will correspond to the timing of the first of the two measures. The date of the HIV-1 RNA virologic failure event will be the first of the two measures; if there is not a confirmatory viral load result then the conservative approach of classifying the single measure has a VF will be used.

KS-IRIS is defined as any IERC-confirmed KS progression that occurs within 12 weeks of initiation of ART that is associated with an increase in CD4+ cell of at least 50 cells/mm^3 above the study entry value and/or decrease in the HIV-1 RNA level by at least 0.5 log below the study entry value prior to or at the time of documented IERC-confirmed KS progression.

Change in KS treatment by week 48 is defined as stopping randomized chemotherapy and initiating a different chemotherapy, regardless of the reason; censored analysis of this outcome will not be done.

4.2.2 Secondary Outcome Methods

<u>Analysis of Events by Week 48</u>

Analyses of events by week 48 will be done in a similar manner to the primary analysis. Kaplan-Meier methods with Greenwood's formula for the variance (see below) will be used to estimate the survival probability at week 48. Participants will be classified as having had an event if they experience the specific component(s) of the secondary outcome within the first 48 weeks on study; for composite secondary outcomes the timing of the first occurrence of the component will be used. Participants who are not classified as an event will be censored at their date of last reported contact (if their total follow up time is less than 48 weeks) or at week 48 (for those with more than 48 weeks of follow up). Time is calculated as the difference from week 0 (step 1 entry) to either event or censor date. The censored analysis will also censor participant follow-up time at the time they start another study chemotherapy (censoring will occur as defined in the previous section).

The absolute difference in 48-week rates between ET+ART and PTX+ART, and BV+ART and PTX+ART will be calculated separately. The 48-week survival probabilities and two-sided 95% CIs for each difference will be determined overall (not adjusted for stratification factors) and with adjustment (weighted) for stratification factors; the inverse of the stratum-specific variance will be used for the stratum weights. Adjustments for stratification factors will be done in a similar manner as the primary analysis. Difference will be calculated for the rates generated from the whole analysis population, and for the censored analysis.

Analysis of KS-IRIS

The analysis of KS-IRIS will be different from the other secondary outcomes of events by week 48. Because there were only a limited number of KS-IRIS events during the course of the study (observed during interim reviews), there will be no formal statistical analyses comparing treatment arms. Instead, descriptive summary listings of all KS-IRIS events will be included in the final analysis report. The date of KS-IRIS will be the date of KS progression reported by the site.

Analyses of Objective Response (OR)

The best overall response on Step 1 for each participant will be determined. The possible categories include: Complete Response (CR), Partial Response (PR), Stable Disease (SD), Progressive Disease (PD), or Not Evaluable (NE). Participants will be classified as NE if they do not have enough follow-up data to determine their KS response. Objective Response is defined as experiencing either a CR or PR (OR = CR+PR). The number and % of participants in each category will be summarized in a table by treatment arm. Logistic regression analysis will be used to compare OR between arms, adjusting for stratification factors (in a similar manner as the primary analysis). The treatment effect from this model will determine if there are higher odds of an OR in ET+ART versus PTX+ART, and in BV+ART versus PTX+ART.

Among those classified with their best response as an OR, the duration of this response will be calculated. Duration will be calculated as the difference in time from first documented response to IERC-confirmed KS progression, death, or change in step. Participants with OR, but without a documented IERC-confirmed KS progression, death, or change in step, will be censored at their last reported date of contact. Kaplan-Meier methods with Greenwood's formula for the variance will be used to determine the median duration of OR for each treatment arm; no formal comparisons between arms will be done.

Time to Event Analysis

Kaplan-Meier methods with Greenwood's formula for the variance will be used to visually display survival curves and estimate median time for each treatment arm. Log-rank test will be used to compare the survival curves between arms; comparisons will be done separately for ET+ART versus PTX+ART and BV+ART versus PTX+ART. Cox-proportional hazards regression models will be adjusted for stratification factors in a similar manner as the primary analysis; hazard ratios (HR), 95% CI for the HR, and p-values will be provided.

Participants will be censored at the time of their last reported date of contact if they did not experience an event. An additional censored analysis will also censor those at the time they start different study chemotherapy (defined in previous section).

Safety Summaries

All safety events will be presented according to MedDRA coding and will be grouped by the most appropriate MedDRA hierarchy at the time of analysis. No formal treatment group comparisons are planned.

Descriptive analysis with tabular summaries will be provided for post-entry Grade 3 or higher signs and symptoms and laboratory toxicities overall and by treatment arm. This will include frequencies and percentages of participants who experienced a post-entry Grade 3 or higher AE; if more than one grade of the same event is reported for a participant then the highest grade will be reported. This summary will be provided for the entire analysis population and will include events with onset dates after randomization to Step 1. A separate summary will also be provided that restricts AEs to the time prior to starting different study chemotherapy (exclude events with onset dates after starting different study chemotherapy).

Descriptive analysis with tabular summaries will be provided in a similar manner for all post-entry treatment-related toxicities and AEs; this analysis will be restricted to the time while on their initial chemotherapy (i.e., will excluded AEs with onset dates after initiate different study chemotherapy). The relationship to treatment for each toxicity or AE is determined by site investigators and is indicated on the CRF as: definitely related, probably related, possibly related, probably not related, or not related. Events that are classified as definitely, probably or possibly related will be considered treatment-related events.

Peripheral Neuropathy Summaries

Assessments of peripheral neuropathy (PN) were recorded on a peripheral neuropathy screening form collected for everyone at screening, week 9, and week 21, and for those on BV or PTX also at weeks 3, 6, 15, and 18. PN is determined by the presence and severity of three components: symptoms, perception of vibrations, and deep tendon reflexes. Each component of PN will be summarized separately and will pool results reported from the right and left side by taking the highest recorded result of the right and left side. Summaries of PN will be by study week and treatment arm and will be done among participants while they were on their initially randomized chemotherapy; frequencies and percentages will be used for this analysis.

Symptomatic PN (SPN) consists of three assessments: (1) pain, aching or burning feet, legs, (2) “pins and needles” in feet, legs, and (3) numbness (lack of feeling) in the feet, legs. SPN that is present is graded on a severity scale from 1 (mild) to 10 (severe). The highest reported severity between the right and left sides across the three assessments will be determined. The presence of symptomatic PN and the severity of the SPN for those with symptoms will be summarized.

Perception of vibration will be summarized as normal or abnormal (mild, moderate, or severe loss). The highest score between the right and left side will be determined. Deep tendon reflexes will be summarized as either: absent/hypoactive, normal, or hyperactive/clonus. The highest reported ankle reflex between the right and left side will be summarized.

Presence of PN is defined as having the all of following results for the three components: (1) presence of symptomatic PN, (2) abnormal perception of vibrations, and (3) absent or hypoactive deep tendon reflexes. The frequency and percent of participants meeting these criteria will also be summarized.

CD4 Count Analysis

Descriptive summaries of change in CD4 cell count will be provided by treatment arm and study week for the analysis population, regardless if treatment was changed. The median and IQR change in CD4 count will be calculated at each study week (12, 24, and 48) and the change will be shown visually in a longitudinal change plot. Sign-ranked tests will be used to determine if CD4 counts changed over time within each treatment group. Formal treatment group comparison will occur at weeks 12 and 48 comparing ET+ART with PTX+ART and BV+ART with PTX+ART using Wilcoxon tests; no adjustment for multiple testing will occur.

ART Adherence and Oral KS

Questionnaires collecting details on ART adherence were administered at multiple study weeks. Summaries of adherence will occur at study weeks 6, 12, 18, 30, and 48. Descriptive summaries of adherence will consist of the frequency and percentage of participants who were adherent, which is defined as either perfect or non-perfect adherence in the last month (defined in a similar manner to A5175 based on participant recall). These summaries will be provided by treatment arm and study week.

The analysis of the prevalence of oral KS at entry will be descriptive and consist of summarizing the frequency and percentage of participant with oral KS lesions at study entry.

Summaries of Study Steps 2-4

Descriptive summaries of secondary outcomes 21-23 will be provided by concurrent treatment arm. The proportion of participants on the given study step who experienced each of the components of the endpoint will be summarized, regardless of prior or subsequent occurrences on other steps. Definitions of each event in the endpoint follow definitions previously noted in prior outcomes. No formal treatment arm comparisons are planned.