STATISTICAL ANALYSIS PLAN

HVTN 119

A phase 1 clinical trial to evaluate the safety and immunogenicity of pDNA vaccines expressing HIV M Group p24Gag conserved elements and/or p55Gag, administered with IL-12 pDNA by intramuscular electroporation, in healthy, HIV-uninfected adult participants

Allan DeCamp, Ph.D.

Erika Rudnicki, M.S.

Version 2.3

17 August 2020

SCHARP
Statistical Center for HIV/AIDS Research and Prevention

APPROVAL SIGNATURE PAGE

HVTN 119

Statistical Analysis Plan

A phase 1 clinical trial to evaluate the safety and immunogenicity of pDNA vaccines expressing HIV M Group p24Gag conserved elements and/or p55Gag, administered with IL-12 pDNA by intramuscular electroporation, in healthy, HIV-uninfected adult participants

I have read this Statistical Analysis Plan and approve its contents.

_____ _____

Allan DeCamp, Ph.D. Date

Staff Scientist

FHCRC, VIDD

_____ _____

Erika Rudnicki, M.S. Date

Statistical Research Associate

FRCRC, SCHARP

SAP Modification History

The version history of, and modifications to, this statistical analysis plan are described below.

SAP Version	Date	Modification
1.0	25 January 2018	--
2.0	07 January 2020	Updated list of SMB tables; added BAMA immunogenicity analysis to section 10.1
2.1	23 January 2020	Added ICS immunogenicity analysis to section 10.1. Added Barnard's test and reverse CDF figure to BAMA immunogenicity analysis.
2.2	06 July 2020	Added lists of tables and figures to be included in the final study report on safety
2.3	04 August 2020	Added ICS COMPASS analysis to section 10.2.

TABLE OF CONTENTS

1. OVERVIEW

This document describes the Statistical Analysis Plan (SAP) for the analysis of interim safety data and immunogenicity data from HVTN 119. As detailed in SCHARP SOP-0013 revision 5 (effective date: 15 August 2016), a SAP is required prior to the first safety monitoring board (SMB) analysis and must be approved by the lead protocol statistician. The plan will be reviewed and updated prior to additional immunogenicity analyses and before the final safety analysis, with all major revisions of the plan archived.

2. PROTOCOL SUMMARY

2.1 TITLE

A phase 1 clinical trial to evaluate the safety and immunogenicity of pDNA vaccines expressing HIV M Group p24Gag conserved elements and/or p55Gag, administered with IL-12 pDNA by intramuscular electroporation, in healthy, HIV-uninfected adult participants

2.2 STUDY PRODUCTS AND ROUTES OF ADMINISTRATION

- p24CE1/2 is a DNA plasmid (pDNA) that encodes 2 cassettes, p24CE1 and p24CE2, each consisting of 7 conserved element (CE) sequences from HIV-1 $p24^{Gag}$ separated by short amino acid spacers. The p24CE1 and p24CE2 proteins differ by 7 amino acids and together cover >98% of HIV-1 Group M sequences worldwide. The p24CE1 cassette is under the control of the human CMV promoter/enhancer and the bovine growth hormone (BGH) polyadenylation signal. The p24CE2 cassette is in the opposite transcriptional orientation and is under the control of the simian CMV promoter and the simian virus 40 (SV40) polyadenylation signal.

- $p55^{gag}$ refers to a DNA plasmid encoding an expression-optimized full length HIV-1 $p55^{Gag}$ protein from the HIV-1 molecular clone HXB2 (Genbank NP_057850).

- *IL-12* refers to GENEVAX® IL-12 DNA Plasmid (Profectus BioSciences, Inc., Tarrytown, New York, USA). The *IL-12* pDNA adjuvant is a dual promoter plasmid that expresses the genes encoding human IL-12 proteins p35 and p40 under separate regulatory control. The p35 subunit is under the control of the human CMV promoter/enhancer and the SV40 polyadenylation signal. The p40 subunit is under the control of the simian CMV promoter and the BGH polyadenylation signal. The plasmid vaccines will be admixed with *IL-12* pDNA to produce a single injectable solution delivered by intramuscular (IM) injection followed by electroporation.
- Placebo: Sodium Chloride for Injection, USP 0.9%.
- Intramuscular TriGrid Delivery System (TDS-IM) (Ichor Medical Systems, San Diego, California), a device for electroporation mediated intramuscular administration of DNA based-biologic candidates.

2.3 SCHEMA

Study arm	Number	Dose (total)	Month 0 (Day 0)	Month 1 (Day 28)	Month 3 (Day 84)	Month 6 (Day 168)
Group 1	25	4 mg 2 mg	p24CE1/2 *IL-12*	p24CE1/2 *IL-12*	p24CE1/2+ p55^{gag} *IL-12*	p24CE1/2+ p55^{gag} *IL-12*
	3	-	Placebo	Placebo	Placebo	Placebo
Group 2	25	4 mg 2 mg	p55^{gag} *IL-12*	p55^{gag} *IL-12*	p55^{gag} *IL-12*	p55^{gag} *IL-12*
	3	-	Placebo	Placebo	Placebo	Placebo
Total	50+ 6 = 56					

Notes

All injections will be delivered IM followed by electroporation (IM/EP). The doses shown will be divided between 2 injection sites, such that 2 mg p24CE1/2 or p55^{gag} pDNA vaccine, or a mixture of 1 mg each, and 1 mg *IL-12* pDNA, or placebo, will be given in each deltoid at every vaccination timepoint.

Enrollment will proceed in both groups simultaneously and will be restricted to US HVTN Clinical Research Sites (CRSs). Across the participating CRSs, enrollment will be restricted to a maximum of 1 participant per day until 10 participants have been enrolled. The HVTN 119 Protocol Safety Review Team (PSRT) will review the cumulative safety data on each of these 10 participants, including at minimum local and systemic reactogenicity data reported for the first 72 hours postvaccination, and will determine whether it is safe to proceed with full enrollment in that group.

2.4 PARTICIPANTS

56 healthy, HIV-uninfected volunteers aged 18 to 50 years; 50 vaccinees, 6 placebo recipients

2.5 DESIGN

Multicenter, randomized, placebo controlled, double-blind trial

2.6 DURATION PER PARTICIPANT

18 months (12 months of scheduled clinic visits (main study) and 1 follow-up contact)

2.7 ESTIMATED TOTAL STUDY DURATION

26 months (includes enrollment, planned safety holds, and follow-up)

3. OBJECTIVES AND ENDPOINTS

3.1 PRIMARY OBJECTIVES AND ENDPOINTS

Primary objective 1:

To evaluate the safety and tolerability of the HIV-1 pDNA vaccines p24CE1/2 and $p55^{gag}$ administered with *IL-12* pDNA by intramuscular injection with electroporation

Primary endpoints 1:

- Frequency and severity of local injection/EP site and systemic reactogenicity signs and symptoms
- Frequency of AEs categorized by MedDRA body system, MedDRA preferred term, severity and assessed relationship to study products. Detailed description of all AEs meeting DAIDS criteria for expedited reporting
- The distribution of values of safety laboratory measures at baseline and at follow-up visits post vaccination
- Number of participants with early discontinuation of vaccinations and reason for discontinuation
- Magnitude of local injection/EP site pain as measured by a visual analog scale
- Distribution of responses to questions regarding acceptability of study injection procedures

Primary objective 2:

To compare the two pDNA prime/boost strategies with respect to breadth (defined as the number of targeted CEs) of $CD4^+$ and $CD8^+$ T-cell responses to conserved regions of the $p24^{Gag}$ protein

Primary endpoint 2:

The number of targeted CEs by $CD4^+$ and $CD8^+$ T-cell responses measured by ICS 2 weeks after the second and fourth vaccinations

3.2 SECONDARY OBJECTIVES AND ENDPOINTS

Secondary objective 1:

To determine the effect of the different prime/boost regimens on the magnitude and polyfunctionality of T-cell responses induced by the different priming vaccinations

Secondary endpoint 1:

Rate, magnitude and polyfunctionality of $CD4^+$ and $CD8^+$ T-cell responses measured by ICS to p24CE and HIV Gag, 2 weeks after the second and fourth vaccinations

Secondary objective 2:

To determine the effect of the different prime/boost regimens on the specificity of T-cell responses induced by the different primes

Secondary endpoint 2:

Epitope specificity of T-cell responses measured by IFN-γ ELISpot to p24CE and HIV Gag, 2 weeks after the second and fourth vaccinations

Secondary objective 3:

To evaluate the humoral immunogenicity of the HIV-1 pDNA vaccines p24CE1/2 and $p55^{gag}$ administered alone, in combination, or sequentially by intramuscular injection with electroporation

Secondary endpoint 3:

Rate and magnitude of humoral immune responses to Gag measured 2 weeks after the 4^{th} vaccination by binding antibody multiplex assay

Secondary objective 4:

To determine the frequency of circulating Tfh in response to each vaccination regimen

Secondary endpoint 4:

Rate and magnitude of $CD4^+$ cTfh measured by ICS

3.3 EXPLORATORY OBJECTIVES

Exploratory objective 1:

To further evaluate immunogenicity of the different vaccine regimens, additional immunogenicity assays may be performed, including on samples from other time points, based on the HVTN Laboratory Assay Algorithm

Exploratory objective 2:

To determine the frequency of plasmablasts in response to each vaccination regimen

Exploratory objective 3:

To evaluate the breadth of the T cell receptor repertoire induced after pDNA vaccination by electroporation

Exploratory objective 4:

To conduct analyses related to furthering the understanding of HIV, immunology, vaccines, and clinical trial conduct

4. COHORT DEFINITION

Recruitment will target enrolling 56 healthy, HIV-uninfected adult participants, aged 18 to 50 years.

To ensure that both men and women will be adequately represented in the trial, the trial will enroll at least approximately 40% of each sex overall. Hence, when approximately 33 participants of one sex are enrolled, recruitment of persons born of that sex will stop.

Since enrollment is concurrent with receiving the first study vaccination, all enrolled participants will provide some safety data. However, for immunogenicity analyses, it is possible that data may be missing for various reasons, such as participants terminating from the study early, problems in shipping specimens, low cell viability of processed peripheral blood mononuclear cells (PBMCs) or high background.

At the enrollment visit, if a participant had the Ichor TDS-IM EP device applied to the participant's arm but the participant was not able to receive any injection of study product, the participant is not considered enrolled into the study, and is referred to as a "device-only participant". Unenrolled, device-only participants will also provide some safety data.

5. POTENTIAL CONFOUNDERS

Characterization of the safety of the vaccine is susceptible to confounding by adverse events not related to the vaccine that by chance occur more often in one arm of the trial than another. Therefore analyses involving adverse events will incorporate the reported relationship to product as assessed by HVTN staff.

6. RANDOMIZATION

The randomization sequence was obtained by computer-generated random numbers and provided to each HVTN CRS through a Web-based randomization system. The randomization was done in blocks to ensure balance across arms. At each institution, the pharmacist with primary responsibility for dispensing study products is charged with maintaining security of the treatment assignments (except in emergency situations as specified in the HVTN MOP).

7. BLINDING

Participants and site staff (except for site pharmacists) will be blinded as to participant treatment arm assignments (i.e., blinded to both group and vaccine/control status). Study product assignments are accessible to those HVTN CRS pharmacists, DAIDS protocol pharmacists and contract monitors, and SDMC staff who are required to know this information in order to ensure proper trial conduct. Any discussion of study product assignment between pharmacy staff and any other HVTN CRS staff is prohibited. The HVTN SMB members also are unblinded to treatment assignment in order to conduct review of trial safety.

When a participant leaves the trial prior to study completion, the participant will be told he or she must wait until all participants are unblinded to learn his or her treatment assignment.

Emergency unblinding decisions will be made by the site investigator. If time permits, the HVTN PSRT should be consulted before emergency unblinding occurs.

8. STATISTICAL ANALYSIS

All data from enrolled participants will be analyzed according to the initial randomization assignment regardless of how many vaccinations they received. In the rare instance that a participant receives the wrong treatment at a specific vaccination time, the Statistical Analysis Plan will address how to analyze the participant's safety data. Analyses are modified intent-to-treat in that individuals who are randomized but not enrolled do not contribute data and hence are excluded. Because of blinding and the brief length of time between randomization and enrollment—typically no more than 4 working days—very few such individuals are expected.

Analyses for primary endpoints will be performed using SAS and R. All other descriptive and inferential statistical analyses will be performed using SAS, StatXact, or R statistical software.

No formal multiple comparison adjustments will be employed for multiple safety endpoints, multiple primary immunogenicity endpoints, or secondary endpoints. However, multiplicity adjustments will be made for certain immunogenicity assays, as discussed below, when the assay endpoint is viewed as a collection of hypotheses (e.g., testing multiple peptide pools to determine a positive response).

8.1 ANALYSIS VARIABLES

The analysis variables consist of baseline participant characteristics, safety, and immunogenicity for primary- and secondary-objective analyses.

8.2 BASELINE COMPARIBILITY

Treatment arms will be compared for baseline participant characteristics using descriptive statistics.

8.3 SAFETY/TOLERABILITY ANALYSIS

Since enrollment is concurrent with receiving the first vaccination, all participants will have received at least 1 vaccination and therefore will provide some safety data.

8.3.1 REACTOGENICITY

The number and percentage of participants experiencing each type of reactogenicity sign or symptom will be tabulated by severity and treatment arm and the percentages displayed graphically by arm. For a given sign or symptom, each participant's reactogenicity will be counted once under the maximum severity for all injection visits. In addition, to the individual types of events, the maximum severity of local pain or tenderness, induration or erythema, and of systemic symptoms will be calculated. Kruskal-Wallis tests will be used to test for differences in severity between arms.

8.3.2 AEs, SAEs, and AESIs

AEs will be summarized using MedDRA System Organ Class and preferred terms. Tables will show by treatment arm the number and percentage of participants experiencing an AE within a System Organ Class or within preferred term category by severity or by relationship to study product. For the calculations in these tables, a participant with multiple AEs within a category will be counted

once under the maximum severity or the strongest recorded causal relationship to study product. Formal statistical testing comparing arms is not planned since interpretation of differences must rely heavily upon clinical judgment.

A listing of SAEs reported to the DAIDS Regulatory Support Center (RSC) Safety Office will provide details of the events including severity, relationship to study product, time between onset and last vaccination, and number of vaccinations received. A separate listing will do the same for AEs of special interest (AESI). AESI for this protocol include but are not limited to autoimmune disorders; a sample list of AESI is provided in Appendix G.

8.3.3 LOCAL LABORATORY VALUES

Box plots of local laboratory values will be generated for baseline values and for values measured during the course of the study by treatment arm and visit. Each box plot will show the first quartile, the median, and the third quartile. Outliers (values outside the box plot) will also be plotted. If appropriate, horizontal lines representing boundaries for abnormal values will be plotted.

For each local laboratory measure, summary statistics will be presented by treatment arm and time point, as well as changes from baseline for post-enrollment values. In addition, the number (percentage) of participants with local laboratory values recorded as meeting Grade 1 AE criteria or above as specified in the Division of AIDS Table for Grading the Severity of Adult and Pediatric Adverse Events will be tabulated by treatment arm for each post-vaccination time point. Reportable clinical laboratory abnormalities without an associated clinical diagnosis will also be included in the tabulation of AEs described above.

8.3.4 REASONS FOR VACCINATION DISCONTINUATION AND EARLY STUDY TERMINATION

The number and percentage of participants who discontinue vaccination and who terminate the study early will be tabulated by reason and treatment arm.

8.4 IMMUNOGENICITY ANALYSIS

8.4.1 GENERAL APPROACH

For the statistical analysis of immunogenicity endpoints, data from enrolled participants will be used according to the initial randomization assignment regardless of how many injections they received. Additional analyses may be performed, limited to participants who received all scheduled injections per protocol. Assay results that are unreliable, from specimens collected outside of the visit window, or from HIV-infected participants post-infection are excluded. Since the exact date of HIV infection is unknown, any assay data from blood draws 4 weeks prior to an infected participant's last seronegative sample and thereafter may be excluded. If an HIV-infected participant does not have a seronegative sample post-enrollment, then all data from that participant may be excluded from the analysis.

Discrete categorical assay endpoints (e.g., response rates) will be analyzed by tabulating the frequency of positive response for each assay by antigen and

treatment arm at each timepoint for which an assessment is performed. Crude response rates will be presented with their corresponding 95% confidence interval estimates calculated using the score test method [1]. Because of the small numbers of control participants in each group, no adjustment will be made to the vaccine arm estimates for the false positive rates in the control arms. Fisher's exact tests will be used to compare the response rates of any 2 vaccine arms, with a significant difference declared if the 2-sided p-value is ≤ 0.05.

For quantitative assay data (e.g., percentage of positive cells from the ICS assay or mean fluorescence intensity from the binding antibody multiplex assay), graphical and tabular summaries of the distributions by antigen, treatment arm, and timepoint will be made. For all primary and secondary immunogenicity endpoints, box plots and plots of estimated reverse cumulative distribution curves will be used for graphical display of all of the study arms. Typically, the results will be shown for each vaccine arm and for the set of control arms pooled into one group.

The differences between treatment groups at 2 weeks post the 2nd and 4th vaccination timepoints will be tested with a nonparametric Wilcoxon rank sum test if the data are not normally distributed and with a 2-sample t-test if the data appear to be normally distributed.

Some immunologic assays have underlying continuous or count-type readouts that are dichotomized into responder/non-responder categories (e.g., ICS and BAMA). If treatment arm differences for these assays are best summarized by a mixture model, then Lachenbruch's test statistic [2] will be used to evaluate the composite null hypothesis of equal response rates in the 2 arms and equal response distributions among responders in the 2 such arms. This test statistic equals the square of a binomial Z-statistic for comparing the response rates plus the square of a Wilcoxon statistic for comparing the response distributions in the subgroup of responders. A permutation procedure is used to obtain a 2-sided p-value. For estimation, differences in response rates between arms will be estimated using the methods described above, and in the subgroup of positive responders, differences in location parameters between arms will be estimated using the methods described above.

More sophisticated analyses employing repeated measures methodology (e.g., linear mixed models or marginal mean models fit by generalized estimating equations) may be utilized to incorporate immune responses over several timepoints and to test for differences over time. However, inference from such analyses would be limited by the small sample size of this study. All statistical tests will be 2-sided and will be considered statistically significant if $p \leq 0.05$.

Based upon previous HVTN trials, missing 17% of immunogenicity results for a specific assay is common due to study participants terminating from the study early, problems in shipping specimens, or low cell viability of processed peripheral blood mononuclear cells (PBMCs). To achieve unbiased statistical estimation and inferences with standard methods applied in a complete-case manner (only including participants with observed data in the analysis), missing data need to be missing completely at random (MCAR). Following the most commonly used definition, MCAR assumes that the probability of an observation being missing does not depend on any participant characteristics (observed or unobserved). When missing data are minimal (specifically if no more than 20% of

participants are missing any values), then standard complete-case methods will be used, because violations of the MCAR assumption will have little impact on the estimates and hypothesis tests.

If a substantial amount of immunogenicity data is missing for an endpoint (at least 1 value missing from more than 20% of participants), then using the methods that require the MCAR assumption may give misleading results. In this situation, analyses of the immunogenicity endpoints at a specific timepoint will be performed using parametric generalized linear models fit by maximum likelihood. These methods provide unbiased estimation and inferences under the parametric modeling assumptions and the assumption that the missing data are missing at random (MAR). MAR assumes that the probability of an observation being missing may depend upon the observed responses and upon observed covariates, but not upon any unobserved factors. Generalized linear models for response rates will use a binomial error distribution and for quantitative endpoints, a normal error distribution. For assessing repeated immunogenicity measurement, linear mixed effects models will be used. If the immunological outcomes are left- and/or right- censored, then the linear mixed effects models of Hughes [3] will be used, because they accommodate the censoring. In addition, secondary analyses of repeated immunogenicity measurements may be done using weighted GEE [4] methods, which are valid under MAR. All of the models described above in this paragraph will include as covariates all available baseline predictors of the missing outcomes.

8.4.2 ANALYSIS OF CD4+ AND CD8+ T-CELL RESPONSE AS MEASURED BY THE ICS ASSAY

The analysis of $CD4^+$ and $CD8^+$ T-cell response rates as measured by the ICS assay will be evaluated and compared as described under the general approach. For each T-cell subset, the positivity call for each peptide pool will include a multiple comparison adjustment for the number of peptide pools used in the assay. In general, the Mixture Models for Single-cell Assays (MIMOSA) statistical framework [5] and/or the Fisher's exact test-based positivity criteria will be used. The magnitude of marginal response will be analyzed as described for quantitative data in the general approach section. For each T-cell subset, graphs will be used to display the background-subtracted magnitudes for each participant by protein, treatment arm and timepoint. Statistical testing comparing the magnitudes will be based on positive responders only. The polyfunctionality of ICS responses will also be analyzed as a secondary endpoint. Besides descriptive plots of the magnitude of polyfunctional responses, the COMPASS (Combinatorial Polyfunctionality analysis of Antigen-Specific T-cell Subsets) statistical framework [6] will also be used to perform joint modelling of multiple T-cell subsets of different cytokine combinations using the functionality score (FS) and the polyfunctionality score (PFS) to summarize the multi-parameter ICS responses.

8.4.3 ANALYIS OF MULTIPLEXED IMMUNOASSAY DATA

The analysis of response rates and response magnitudes will be evaluated and compared as described under the general approach.

8.5 ANALYSIS PRIOR TO END OF SCHEDULED FOLLOW-UP VISITS

Any analyses conducted prior to the end of the scheduled follow-up visits should not compromise the integrity of the trial in terms of participant retention or safety or immunogenicity endpoint assessments. In particular, early unblinded analyses by treatment assignment require careful consideration and should be made available on a need to know basis only.

8.5.1 SAFETY

During the course of the trial, unblinded analyses of safety data will be prepared approximately every 4 months during the main study for review by the SMB. Ad hoc safety reports may also be prepared for SMB review at the request of the HVTN PSRT. The HVTN leadership must approve any other requests for unblinded safety data prior to the end of the scheduled follow-up visits.

8.5.2 IMMUNOGENICITY

An unblinded statistical analysis by treatment assignment of a primary immunogenicity endpoint may be performed when all participants have completed the corresponding primary immunogenicity visit and data are available for analysis from at least 80% of these participants. Similarly, an unblinded statistical analysis by treatment assignment of a secondary or exploratory immunogenicity endpoint may be performed when all participants have completed the corresponding immunogenicity visit and data are available for analysis from at least 80% of these participants. However, such analyses for a secondary or exploratory immunogenicity endpoint will only take place after at least one of the primary immunogenicity endpoints of the same class (humoral or cell-mediated) or, if no primary endpoint of the same class, at least one of the primary immunogenicity endpoints reaches the aforementioned threshold. The Laboratory Program will review the analysis report prior to distribution to the protocol chairs, DAIDS, vaccine developer, and other key HVTN members and investigators. Distribution of reports will be limited to those with a need to know for the purpose of informing future trial-related decisions. The HVTN leadership must approve any other requests for HVTN immunogenicity analyses prior to the end of the scheduled follow-up visits.

9. SAFETY TABLES AND FIGURES

9.1 LIST OF SMB TABLES

- Enrollment Report
- Demographics and Study Product Administration Frequencies
- Overall Protocol Status
- Discontinuation Status Listing
- Study Product Administration Errors Listing
- Maximum Local Reactogenicity Summary
- Maximum Systemic Reactogenicity Summary
- Severe or Life-Threatening Local Reactogenicities Listing

- Severe or Life-Threatening Systemic Reactogenicities Listing
- Moderate, Severe, or Life-Threatening Erythema/Induration Listing
- AEs by Body System and Severity
- AEs by Preferred Term and Severity – Includes Severe, Life-Threatening, or Fatal Events Only
- AEs by Preferred Term and Severity – Includes All Severities
- AEs by Preferred Term and Severity – Includes Related Events Only
- AEs by Preferred Term and Severity – Includes Events Related to the EP Device Only
- AEs by Preferred Term and Relationship to Study Product – Includes Events of Any Relationship
- Severe, Life-Threatening, or Fatal Adverse Events Listing
- Adverse Events of Special Interest Listing
- Related Adverse Events Listing
- Device Related Adverse Events Listing
- EAEs Reported to the RSC Listing
- Pregnancy Listing
- HIV Infection Listing
- Listing of Device-Only Participants
- Listing of Device-Only Participant AEs

9.2 LIST OF SMB FIGURES

- Maximum Local Reactogenicities
- Maximum Systemic Reactogenicities
- Boxplots for platelets, WBC, lymphocytes, neutrophils, hemoglobin, alkaline phosphate, AST, ALT, and creatinine at baseline and each post-vaccination follow-up visit

9.3 LIST OF SAFETY FSR TABLES

- Enrollment Report
- Demographics and Study Product Administration Frequencies
- Overall Protocol Status
- Discontinuation Status Listing
- Study Product Administration Errors Listing
- Maximum Local Reactogenicity Summary
- Maximum Systemic Reactogenicity Summary
- Severe or Life-Threatening Local Reactogenicities Listing
- Severe or Life-Threatening Systemic Reactogenicities Listing
- Moderate, Severe, or Life-Threatening Erythema/Induration Listing
- AEs by Body System and Severity
- AEs by Preferred Term and Severity – Includes Severe, Life-Threatening, or Fatal Events Only
- AEs by Preferred Term and Severity – Includes All Severities
- AEs by Preferred Term and Severity – Includes Related Events Only
- AEs by Preferred Term and Severity – Includes Events Related to the EP Device Only

- AEs by Preferred Term and Relationship to Study Product – Includes Events of Any Relationship
- Severe, Life-Threatening, or Fatal Adverse Events Listing
- Adverse Events of Special Interest Listing
- Related Adverse Events Listing
- Device Related Adverse Events Listing
- EAEs Reported to the RSC Listing
- Pregnancy Listing
- HIV Infection Listing
- Listing of Device-Only Participants
- Listing of Device-Only Participant AEs
- Social Impact Summary
- End of Study Diagnostic Testing Results for Vaccinees
- EP Acceptability Questionnaire Results
- Local Lab Values Meeting Grade 1 AE Criteria or Above
- Local Lab Value Summary Statistics Tables

9.4 LIST OF SAFETY FSR FIGURES

- Maximum Local Reactogenicities
- Maximum Systemic Reactogenicities
- Boxplots for platelets, WBC, lymphocytes, neutrophils, hemoglobin, alkaline phosphate, AST, ALT, and creatinine at baseline and each post-vaccination follow-up visit
- Boxplots with visual analog scale (VAS) pain assessment scores at each vaccination visit

10. ASSAY-SPECIFIC TABLES AND FIGURES FOR PROTOCOL TEAM REPORTS

10.1 Binding Antibody Multiplex Assay (BAMA)

Serum HIV-1-specific IgG antibody responses are measured at the 1:50 dilution against the $p24^{gag}$ antigen on all available samples at visit 2 (month 0/baseline) and visit 11 (Month 6.5/2 weeks post 4th vaccination) (Tomaras and Yates et al., J Virology 2008). The assay readout is from a Bio-Plex instrument (Bio-Rad). The Bioplex software provides 2 readouts: a background-subtracted mean fluorescent intensity (MFI), where background refers to a plate level control (i.e., a blank well run on each plate), and a concentration based on a standard curve. Net MFI (MFI*) is MFI minus Neg, where 'Neg' refers to a sample-specific background measure. Samples from post-enrollment visits are declared to have positive responses if they meet three conditions: (1) the MFI* values are greater than the antigen-specific cutoff (based on the 95th percentile of the baseline visit serum samples and at least 100 MFI), (2) the MFI* values are greater than 3 times the baseline (day 0) MFI* values, and (3) the MFI values are greater than 3 times the baseline MFI values. Net MFI is used to summarize the magnitude at a given time-point. If sufficient immunogenicity is observed, samples may be

titrated to calculate antibody titers. Titrations are quantified by EC50/AUC. If titrations are performed, the geometric mean titer will be compared between the treatment groups. The overall difference in geometric mean titer between groups at each time point will also be compared.

Several criteria are used to determine if data from an assay are acceptable and can be statistically analyzed. The blood draw date must be within the allowable visit window as determined by the protocol. Second, if the blank bead negative control exceeds 5,000 MFI, the sample will be repeated. If the repeat value exceeds 5,000 MFI, the sample will be excluded from analysis due to high background.

The following comparisons of the immune response rates and magnitude (MFI* values (overall and among positive responders only)) will be done:

- Fisher's exact test and Barnard's test are used to compare response rates between pairs of groups, and the Wilcoxon rank sum test is used to compare response magnitudes (net MFI) among positive responders between pairs of groups. Group comparisons include T1 vs. control (pooled), T2 vs. control (pooled), and T1 vs. T2. P-values are two-sided and unadjusted for multiple comparisons.

Contingent upon these results, assays may be performed on additional timepoints, isotypes, and/or antigens.

List of Tables

Tables are stratified by lab, isotype, antigen, visit number, visit month, and treatment group.

- Response rate table by isotype, antigen, visit, and treatment group. Positivity criteria will be stated in footnote.
- Summary statistics (minimum, 25th percentile, median, 75th percentile, maximum, mean, standard deviation) among all participants, by treatment group.
- Summary statistics (minimum, 25th percentile, median, 75th percentile, maximum, mean, standard deviation) among positive responders, by treatment group.
- Listing of magnitudes (MFI* values) among positive responders.
- Response rate comparisons.
- Response magnitude comparisons among positive responders, if there are at least 5 responders within each group at a given visit.

List of Figures

Figures are stratified by lab, isotype, antigen, visit number, visit month, and treatment group.

- Boxplots of MFI* values by isotype, antigen, visit, and treatment group. Individual-level data for post vaccination visits will be connected by lines, with treatment group side by side, one isotype and antigen per graph.
- Reverse CDFs of MFI* by isotype, antigen, visit, and treatment group.

- If titrations are measured, boxplots of binding titration (EC50/AUC) by isotype, antigen, visit and group/treatment arm. Individual-level data for post vaccination visits will be connected by lines, with treatment group side by side, one isotype and antigen per graph.

10.2 Intracellular Cytokine Staining Assay (ICS)

HIV-1-specific CD4+ and CD8+ T-cell responses are measured by intracellular cytokine staining as previously described (De Rosa, 2012 & Horton, 2007). A validated 28-color panel (experiment Assay ID 114; analysis plan 048) is used to evaluate Gag CE total and Gag p55 peptides at visit 6 (month 1.5/2 weeks post 2nd vaccination) and visit 11 (Month 6.5/2 weeks post 4th vaccination). CD4+ and CD8+ T cells expressing IL-2 and/or IFN-γ are included in the primary analysis. Marginal expression of CD40L for CD4+ T cells are included as a secondary analysis.

Several criteria are used to determine if data from an assay were acceptable and could be statistically analyzed. The blood draw date must have been within the allowable visit window as determined by the protocol. Post-infection samples from HIV-infected participants are excluded. After sample thawing and overnight incubation, the viability of the PBMC must have been 66% or greater for testing to have proceeded. If it was not, a new specimen for that participant at that time point was thawed for testing. If the PBMC viability of the second thawed aliquot was below this threshold, the ICS assay was not performed and no data were reported to the statistical center for the participant-time point. For the negative control acceptance criteria, if the average cytokine response for the negative control wells was above 0.1% for either the CD4+ or CD8+ T cells, the sample was retested. If the retested results were above 0.1%, the data were excluded from analysis; otherwise, the retest data were used.

The total numbers of CD4+ and CD8+ T cells must also have exceeded certain thresholds. If the number of CD4+ T cells was < 10,000 or the number of CD8+ T cells was < 5,000 for any of the HIV-1 peptide pools or for one of the negative control replicates for a particular sample, data for that stimulation were filtered. If both negative control replicates failed for number of T cells, the sample was retested. If one negative control replicate failed for number of T cells, the negative control replicate with sufficient cells was used.

To assess positivity for a peptide pool within a T-cell subset, a two-by-two contingency table was constructed comparing the HIV-1 peptide stimulated and negative control data. The four entries in each table were the number of cells positive for IL-2 and/or IFN-γ and the number of cells negative for IL-2 and IFN- γ for both the stimulated and the negative control data. If both negative control replicates were included, then the average number of total cells and the average number of positive cells were used. A one-sided Fisher's exact test was applied to the table, testing whether the number of cytokine-producing cells for the

stimulated data was equal to that for the negative control data. Since multiple individual tests (for each peptide pool) were conducted simultaneously, a multiplicity adjustment was made to the individual peptide pool p-values using the Bonferroni-Holm adjustment method. If the adjusted p-value for a peptide pool was ≤ 0.00001, the response to the peptide pool for the T-cell subset was considered positive. Because the sample sizes (i.e., total cell counts for the T-cell subset) were large, e.g., as high as 100,000 cells, the Fisher's exact test has high power to reject the null hypothesis for very small differences. Therefore, the adjusted p-value significance threshold was chosen stringently (≤0.00001). If at least one peptide pool for a specific HIV-1 protein was positive, then the overall response to the protein was considered positive. If any peptide pool was positive for a T-cell subset, then the overall response for that T-cell subset was considered positive.

The following comparisons of the immune response rates and magnitudes (among positive responders only) will be done:

- Fisher's exact test and Barnard's test are used to compare response rates between pairs of groups, and the Wilcoxon rank sum test is used to compare response magnitudes among positive responders between pairs of groups. Group comparisons include T1 vs. control (pooled), T2 vs. control (pooled), and T1 vs. T2. P-values are two-sided and unadjusted for multiple comparisons.

Contingent upon these results, additional assays and/or analyses may be performed.

The following comparisons of FS and PFS will be done:

- Wilcoxon rank sum tests are used to compare FS and PFS between pairs of groups at each visit. Group comparisons include T1 vs. control (pooled), T2 vs. control (pooled), and T1 vs. T2. P-values are two-sided and adjusted for multiple antigens and visits via the false-discovery rate (FDR) method. Comparison results are considered statistically significant if p-value < 0.05 and the FDR-adjusted q-value < 0.2.

List of Tables

Tables are stratified by T cell subset, cytokine, antigen, visit number (visit month), and treatment group.

- Response rate table by T cell subset, cytokine, antigen, visit number (visit month), and treatment group.
- Summary statistics (minimum, 25th percentile, median, 75th percentile, maximum, mean, standard deviation) among all participants, by T cell subset, cytokine, antigen, visit number (visit month), and treatment group.
- Summary statistics (minimum, 25th percentile, median, 75th percentile, maximum, mean, standard deviation) among positive responders, by T cell subset, cytokine, antigen, visit number (visit month), and treatment group.

- Listing of magnitudes among positive responders.
- Response rate comparisons.
- Response magnitude comparisons among positive responders, if there are at least 5 responders within each group at a given visit.
- Summary statistics (minimum, 25th percentile, median, 75th percentile, maximum, mean, standard deviation) of FS and PFS by T-cell subset, HIV protein, visit, and treatment group.
- FS and PFS comparison between treatment groups by T-cell subset, HIV protein, and visit.

List of Figures

Figures are stratified by T cell subset, cytokine, antigen, visit number (visit month), and treatment group.

- Boxplots of background-subtracted percent of cells expressing IL-2 and/or IFN-γ by T cell subset, antigen, visit number (visit month), and treatment group.
- Spaghetti plots of background-subtracted percent of cells expressing IL-2 and/or IFN-γ over time, by T cell subset, antigen and treatment group.
- Reverse CDFs of background-subtracted percent of cells expressing IL-2 and/or IFN-γ and/or CD40L by T cell subset, antigen and treatment group.
- Individual heatmaps showing posterior probabilities of antigen-specific responses from COMPASS by T-cell subset, HIV protein, and visit.
- Summary heatmaps showing posterior probabilities of antigen-specific responses from COMPASS by T-cell subset, HIV protein, and visit.
- Boxplots of FS/PFS by treatment group, T-cell subset, HIV protein, and visit.

11. REFERENCES

1. Agresti A, Coull BA. Approximate is better than "exact" for interval estimation of binomial proportions. Am Stat 1998;52:119-26.
2. Lachenbruch PA. Comparisons of two-part models with competitors. Stat Med 2001;20:1215-34.
3. Hughes JP. Mixed effects models with censored data with application to HIV RNA levels. Biometrics 1999;55:625-9.
4. Rotnitzky A, Robins J. Analysis of semi-parametric regression models with non-ignorable non-response. Stat Med 1997;16:81-102.
5. Finak G, McDavid A, Chattopadhyay P, Dominguez M, De RS, Roederer M, Gottardo R. Mixture models for single-cell assays with applications to vaccine studies. Biostatistics 2014;15:87-101.
6. Lin L, Finak G, Ushey K, Seshadri C, Hawn TR, Frahm N, et al. COMPASS identifies T-cell subsets correlated with clinical outcomes. Nat Biotechnol. 2015;33(6):610-6.