STATISTICAL ANALYSIS PLAN FOR SAFETY AND IMMUNOGENICITY

Protocol HVTN 122

A phase 1 double-blind, randomized clinical trial to evaluate the safety and immunogenicity of a recombinant oligomeric gp145 clade C Env protein (gp145 C.6980) in healthy, HIV-1–uninfected adult participants in the US

Date: 28 October 2019

SAP version: 1.3

Prepared by

Protocol Statisticians Yunda Huang, PhD
Hua Zheng, M.S.

Approval Signature Page

HVTN 122

Statistical Analysis Plan for Safety and Immunogenicity

A phase 1 double-blind, randomized clinical trial to evaluate the safety and immunogenicity of a recombinant oligomeric gp145 clade C Env protein (gp145 C.6980) in healthy, HIV-1–uninfected adult participants in the US

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

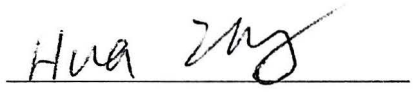
11/08/2019

Yunda Huang, PhD — Date

Principle Staff Scientist

Vaccine and Infectious Disease Division

Fred Hutchinson Cancer Research Center



08NOV2019

Hua Zheng, MS — Date

Statistical Research Associate

Fred Hutchinson Cancer Research Center, SCHARP

If applicable, include other signatures: NA

SAP Modification History

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

Date: 10 July 2018
SAP version: Version *1.0*
Modifications: First draft concerning only the analysis of safety endpoints.

Date: 20 August 2019
SAP version: Version *1.1*
Modifications:

- Changed study SRA to Hua Zheng due to the departure of the previous study SRA.
- Added immunogenicity analysis (ICS Assay) in Section 10.5.3.

Date: 24 September 2019
SAP version: Version *1.2*
Modifications:

- Added immunogenicity analysis (BAMA assay) in Section 10.5.4, per the BAMA assay study plan.
- Added listing of tables and figures of BAMA assay data in Section 12.2.

Date: 28 October 2019
SAP version: Version *1.3*
Modifications:

- Added immunogenicity analysis (neutralizing antibody assay) in Section 10.5.5, per the neutralizing antibody assay study plan.
- Added listing of tables and figures of neutralizing antibody assay data in Section 12.3.

Table of Contents

1 LIST OF ABBREVIATIONS AND ACRONYMS

AE	Adverse Experience
EAE	Expedited Adverse Experience
FSR	Final Study Report
PT	Protocol Team
RSC	Regulatory Support Center
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SMB	Safety Monitoring Board

2 OVERVIEW

The following describes the Statistical Analysis Plan (SAP) for the analysis of safety and immunogenicity data from HVTN 122 for Safety Monitoring Board (SMB) and Protocol Team immunogenicity data reports. As detailed in SCHARP SOP-0013, Revision 5 (effective date: August 15, 2016), this SAP is required prior to the first analysis of each respective data type and must be approved by the lead protocol statistician. The plan will be reviewed and updated prior to the first SMB report and the first PT report of each immunogenicity data type with all major revisions of the plan archived.

3 PROTOCOL SUMMARY

Title

A phase 1 double-blind, randomized clinical trial to evaluate the safety and immunogenicity of a recombinant oligomeric gp145 clade C Env protein (gp145 C.6980) in healthy, HIV-1–uninfected adult participants in the US

Study products and routes of administration

- **gp145 C.6980**: gp145 C.6980 is an HIV-1 subtype C recombinant gp145 Env protein produced in CHO (Chinese hamster ovary) cells and derived from an acute Tanzanian clade C Env (C06980.v0.c22). gp145 C.6980 will be given with **aluminum hydroxide adjuvant**, mixed together at the clinical site. 300mcg (Group 1) and 100mcg (Group 2) of gp145 C.6980 will be admixed with **aluminum hydroxide adjuvant** containing ~600mcg ionic aluminum. The resulting vaccine/adjuvant mixtures will be given as single 1 mL intramuscular (IM) injections into the deltoid.

- **Placebo**: Sodium Chloride for Injection, 0.9% delivered as a 1 mL IM injection

Schema

Study arm	Number	Dose	Month 0 (Day 0)	Month 2 (Day 56)	Month 6 (Day 168)
Group 1	25	300 mcg	gp145 C.6980 + alum	gp145 C.6980 + alum	gp145 C.6980 + alum
Group 2	15	100 mcg	gp145 C.6980 + alum	gp145 C.6980 + alum	gp145 C.6980 + alum
Group 3	5	NA	placebo	placebo	placebo
Total	45 (40/5)				

Note

Enrollment will proceed in all study arms simultaneously and will be restricted to a maximum of 1 participant per group per day until 20% of participants (ie, the first 9 participants [5 in Group 1, 3 in Group 2, and 1 in Group 3]) have been enrolled. The PSRT will review the safety and reactogenicity data reported for the first 2 weeks following the first vaccination for each of these participants and will determine whether it is safe to proceed with full enrollment.

Participants

45 healthy, HIV-1-uninfected volunteers aged 18 to 50 years; 40 vaccinees, 5 placebo recipients in the US

Design

Multicenter, randomized, placebo-controlled, double-blind trial

Duration per participant

12 months of scheduled clinic visits

Estimated total study duration

17 months (includes enrollment, planned safety hold, and follow-up)

Investigational New Drug (IND) sponsor

DAIDS, NIAID, NIH, DHHS (Bethesda, Maryland, USA)

Study product providers

- **gp145 C.6980**: Division of AIDS, NIAID, NIH (Bethesda, Maryland, USA)
- **Aluminum hydroxide adjuvant**: Vaccine Research Center, NIAID, NIH (Frederick, Maryland, USA)

Core operations

HVTN Vaccine Leadership Group/Core Operations Center, Fred Hutchinson Cancer Research Center (FHCRC) (Seattle, Washington, USA)

Statistical and data management center (SDMC)

Statistical Center for HIV/AIDS Research and Prevention (SCHARP), FHCRC (Seattle, Washington, USA)

HIV diagnostic laboratory

University of Washington Virology Specialty Laboratory (UW-VSL) (Seattle, Washington, USA)

Endpoint assay laboratories

- Duke University Medical Center (Durham, North Carolina, USA)
- FHCRC/University of Washington (Seattle, Washington, USA)

Study sites

HVTN Clinical Research Sites (HVTN CRSs) to be specified in the Site Announcement Memo

Safety monitoring

HVTN 122 PSRT; HVTN Safety Monitoring Board (SMB)

4 OBJECTIVES AND ENDPOINTS

4.1 Primary objectives and endpoints:

Primary objective 1

- To evaluate the safety and tolerability of 1 injection of gp145 C.6980 at 2 dose levels with alum adjuvant in HIV-seronegative low risk adults

Primary endpoints 1

- Frequency of severe local and systemic reactogenicity signs and symptoms: pain, tenderness, maximum severity of pain and/or tenderness, erythema, induration, fever, malaise/fatigue, myalgia, headache, nausea, vomiting, chills, arthralgia, maximum severity of systemic symptoms (graded using the Division of AIDS Table for Grading the Severity of Adult and Pediatric Adverse Events. Corrected Version 2.1, dated July 2017)
- Frequency of AEs by treatment arm, by body system, Medical Dictionary for Regulatory Activities (MedDRA) preferred term, severity, and assessed relationship to study products (graded using the Division of AIDS Table for Grading the Severity of Adult and Pediatric Adverse Events. Corrected Version 2.1, dated July 2017)
- Serious adverse events (SAEs) throughout the active surveillance period
- Laboratory measures of safety: white blood cells (WBC), neutrophils, lymphocytes, hemoglobin, platelets, ALT, and creatinine at baseline and following vaccinations, by treatment arm (graded using the Division of AIDS Table for Grading the Severity of Adult and Pediatric Adverse Events. Corrected Version 2.1, dated July 2017)

4.2 Secondary objectives and endpoints:

Secondary objective 1

- To evaluate the safety and tolerability of 3 injections of gp145 C.6980 at 2 dose levels with alum adjuvant in HIV-seronegative low risk adults

Secondary endpoints 1

- Frequency of severe local and systemic reactogenicity signs and symptoms: pain, tenderness, maximum severity of pain and/or tenderness, erythema, induration, fever, malaise/fatigue, myalgia, headache, nausea, vomiting, chills, arthralgia, maximum severity of systemic symptoms (graded using the Division of AIDS Table for Grading the Severity of Adult and Pediatric Adverse Events. Corrected Version 2.1, dated July 2017)

- Frequency of AEs by treatment arm, by body system, Medical Dictionary for Regulatory Activities (MedDRA) preferred term, severity, and assessed relationship to study products (graded using the Division of AIDS Table for Grading the Severity of Adult and Pediatric Adverse Events. Corrected Version 2.1, dated July 2017)

- SAEs throughout the active surveillance period

- Laboratory measures of safety: white blood cells (WBC), neutrophils, lymphocytes, hemoglobin, platelets, ALT, and creatinine at baseline and following vaccinations, by treatment arm (graded using the Division of AIDS Table for Grading the Severity of Adult and Pediatric Adverse Events. Corrected Version 2.1, dated July 2017)

Secondary objective 2

- To characterize the immunogenicity of 1 injection of gp145 C.6980 at 2 dose levels with alum adjuvant in HIV-seronegative low risk adults

Secondary endpoints 2

- Response rates and levels of vaccine-induced binding antibodies to HIV proteins measured by the binding antibody multiplex assay (BAMA) at 2 weeks following the first vaccination

Secondary objective 3

- To characterize the immunogenicity of 3 injections of gp145 C.6980 at 2 dose levels with alum adjuvant in HIV-seronegative low risk adults

Secondary endpoints 3

- Response rates and levels of vaccine-induced binding antibodies to HIV proteins measured by the BAMA at 2 weeks following the third vaccination

- Response rates and levels of CD4+ and CD8+ T cells measured by intracellular cytokine staining (ICS) at 2 weeks following the third vaccination

- Response rates and levels of neutralizing antibody responses against HIV-1 isolates at 2 weeks following the third vaccination

4.3 Exploratory objectives:

Exploratory objective 1

- To describe participants' baseline characteristics (eg, previous exposure to vaccinia) and their impact on vaccine-induced immune responses

Exploratory objective 2

- To determine the frequency of circulating Tfh, Tfr, and plasmablasts in response to each vaccination regimen

Exploratory objective 3

- To further evaluate immunogenicity of each vaccine regimen, additional immunogenicity assays may be performed, including on samples from other timepoints, based on the HVTN Laboratory Assay Algorithm.

Exploratory objective 4

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

5 COHORT DEFINITION

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 participant's safety data may also be excluded from the randomized group and examined separately. 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.

6 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.

7 RANDOMIZATION

The randomization sequence will be obtained by computer-generated random numbers and provided to each HVTN CRS through FSTRF. The randomization will be 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 Manual of Operations (MOP).

8 BLINDING

Participants and site staff (except for site pharmacists) will be blinded as to participant treatment arm assignments (eg, vaccine or control). 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 122 Protocol Safety Review Team (PSRT) should be consulted before emergency unblinding occurs.

9 SAMPLE SIZE

The goal of the safety evaluation for this study is to identify safety concerns associated with product administration. The ability of the study to detect SAEs can be expressed by the true event rate above which at least 1 SAE would likely be observed and the true event rate below which no events would likely be observed. Specifically, for each vaccine arm of the study (n = 15, 25), there is a 90% chance of observing at least 1 event if the true rate of such an event is 14.3% or 8.9% or more, respectively; and there is a 90% chance of observing no events if the true rate is 0.4% or 0.7% less, respectively. For vaccine arms combined (n = 40), there is a 90% chance of observing at least 1 event if the true rate of such an event is 5.6% or more; and there is a 90% chance of observing no events if the true rate is 0.26% or less. As a reference, in previous AVEG HIV vaccine trials, 3.5% of control participants experienced an SAE; in HVTN vaccine trials from December 2000 through April 2014, about 4% of participants who received placebos experienced an SAE.

Probabilities of observing 0 and 2 or more events among 15, 25, or 40 vaccine recipients are presented in Table 1 (same as Table 6-1 in the Protocol) below for a range of possible true AE rates. These calculations provide a complete picture of the sensitivity of this study design to identify potential safety problems with the vaccine. For example, we see that if the true rate of AEs is 1%, there is an 86%, 78%, or 67% chance that no events will be observed in the vaccine arm of 15, 25, or 40 participants, respectively.

Table 1: Probability of observing 0 events, 1 or more events, and 2 or more events, among n=15, 25 or 40 vaccine recipients, for different true event rates

True event rate (%)	Pr(0/15)	Pr(2+/15)	Pr(0/25)	Pr(2+/25)	Pr(0/40)	Pr(2+/40)
1	0.860	0.010	0.778	0.026	0.669	0.061
3.5	0.586	0.095	0.410	0.218	0.240	0.411
5	0.463	0.171	0.277	0.358	0.129	0.601
10	0.206	0.451	0.072	0.729	0.015	0.920
20	0.035	0.833	0.004	0.973	<0.001	0.999
30	0.005	0.965	<0.001	0.998	<0.001	>0.999
40	<0.001	0.995	<0.001	>0.999	<0.001	<0.999

An alternative way of describing the statistical properties of the study design is in terms of the 95% confidence interval (CI) for the true rate of an AE based on the observed data. Table 2 (same as Table 6-2 in the Protocol) shows the 2-sided 95% CIs for the probability of an event based on a particular observed rate. Calculations are done using the score test method [36]. If none of the 15, 25 or 40 vaccine recipients experience a safety event, the 95% 2-sided upper confidence bound for the true rate of such events is 21.8%, 13.7% or 8.81%, respectively.

Table 2: Two-sided 95% confidence intervals based on observing a particular rate of safety endpoints for arms of size n_1 and n_2

Observed event rate	CI (%)
0/15	(0.00, 21.8)
1/15	(0.17, 31.9)
2/15	(1.66, 40.5)
0/25	(0.00, 13.7)
1/25	(0.10, 20.4)
2/25	(0.98, 26.0)
0/40	(0.00, 8.81)
1/40	(0.06, 13.2)
2/40	(0.61, 16.9)

10 STATISTICAL ANALYSIS

All data from enrolled participants will be analyzed according to the initial randomization assignment regardless of how many vaccinations they received.

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 (eg, testing multiple peptide pools to determine a positive response).

10.1 Analysis variables

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

10.2 Baseline comparability

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

10.3 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.

10.3.1 Reactogenicity

The number and percentage of participants experiencing each type of reactogenicity sign or symptom will be tabulated by severity and treatment group and the percentages displayed graphically by group. 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. For the final study analysis, Kruskal-Wallis tests will be used to test for differences in severity between groups.

10.3.2 AEs and SAEs

AEs will be summarized using MedDRA System Organ Class and preferred terms. Tables will show by treatment group 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 groups 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.

10.3.3 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 group 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 timepoint, as well as changes from baseline for postenrollment 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 (section 9.8 of the Protocol) will be tabulated by treatment arm for each postvaccination timepoint. Reportable clinical laboratory abnormalities without an associated clinical diagnosis will also be included in the tabulation of AEs described above.

10.3.4 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 group.

10.4 Analyses 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.

10.4.1 Safety

During the course of the trial, unblinded analyses of safety data will be prepared approximately every 4 months for review by the SMB. Ad hoc safety reports may also be prepared for SMB review at the request of the HVTN 122 PSRT. The HVTN leadership must approve any other requests for unblinded safety data prior to the end of the scheduled follow-up visits.

10.4.2 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 1 of the primary immunogenicity endpoints of the same class (humoral or cell-mediated) or, if no primary endpoint of the same class, at least 1 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.

10.5 Immunogenicity analysis

10.5.1 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 postinfection 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 postenrollment, then all data from that participant may be excluded from the analysis.

Discrete categorical assay endpoints (eg, 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, no adjustment will be made to the vaccine arm estimates for the false positive rates in the control arms. Barnard or Fisher's exact tests, as specified in the SAP, will be used to compare the response rates of the 2 vaccine arms, with a significant difference declared if the 2-sided p-value is ≤ 0.05. In general Barnard's is preferred since under most circumstances it is more powerful than Fisher's [2].

For quantitative assay data (eg, IgG binding Ab response from the multiplex assay or CD4+/CD8+ T cell response from the ICS 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 control arm separately.

The difference between arms at a specific timepoint may 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. An appropriate data transformation (eg, log10 transformation) may be applied to better satisfy assumptions of symmetry and homoscedasticity (constant variance).

10.5.2 Multivariate display of immunogenicity endpoints

Data visualization techniques may be used to explore the relationship among immunogenicity readouts. The set of readouts may be based on 1 of the primary endpoints (e.g., ICS), on the set of primary endpoints, or on immunogenicity endpoints that also include secondary or exploratory

endpoints. To understand the relationship between pairs of readouts, scatter plots may be used when the number of readouts is small or for a larger number of readouts, a heat map showing the degree of correlation between any 2 pairs. Principal component analysis (PCA) and associated "biplots" of the scores and loadings are particularly useful to understand associations between readouts, especially when readouts are correlated [3]. PCA is a method to reduce the dimensionality of the number of readouts to a smaller set of values (principal components) that are normalized linear combinations of the readouts in such a way that the first principal component accounts for the most variability in the data and subsequent components, while maximizing variability, are uncorrelated with each other. A 'biplot' displays the first and second principal component scores and principal component loadings. The x-axis is the value from the first principal component and the y-axis is the second principal component, where each axis label includes the percentage of variation in the total set of readouts captured by the principal component. The top axis is the first principal component loadings and the right axis is the second principal component loadings. An arrow is drawn for each immunogenicity readout (eg, Env-specific CD4+ T cell polyfunctionality score, Env-specific CD8+ T cell total magnitude) from the origin to the point defined by its first 2 principal component loadings. The length of the arrow represents the amount of total variation of the set of readouts captured by the given readout. The direction of an arrow conveys the extent to which the variation of a readout is in the direction of the first or second principal component. The angle between 2 arrows conveys information about the correlation of the 2 readouts, with a 0 degree angle denoting perfect correlation and a 90 degree angle denoting no correlation. Each arrow on the biplot is labeled by the immunogenicity readout it represents. A biplot is annotated with key meta-information such as the treatment arm (most common application) or a demographic category. Depending on the application, K-means clustering and hierarchical clustering may also be applied for multivariate graphical display of immunogenicity readouts.

10.5.3 Analysis of CD4+ and CD8+ T-cell response as measured by the ICS assay

The frequency and magnitude of CD4+ and CD8+ T-cells secreting IFN-γ and/or IL-2 as measured by the 17-color ICS assay will be evaluated and compared as described under the general approach in Section 10.5.1. For each T-cell subset, the Fisher's exact test-based positivity criteria will be used to determine the positivity call for responses to each peptide pool. The Mixture Models for Single-cell Assays (MIMOSA) statistical framework [4] for positivity may also be used as supporting analyses. If requested, the magnitude of marginal response may 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. Given that multiple cytokines are measured in this study, the polyfunctionality of ICS responses may also be analyzed as an exploratory endpoint when there are a sufficient number of participants with a positive CD4+ or CD8+ IFN-γ and/or IL-2 response. To this end, besides descriptive plots of the magnitude of polyfunctional responses, the Combinatorial Polyfunctionality analysis of Antigen-Specific T-cell Subsets (COMPASS) statistical framework [5] may also be used to perform joint modelling of multiple T-cell subsets of different cytokine combinations. For example, the functionality score (FS) and the polyfunctionality score (PFS) may be used to summarize the multi-parameter ICS responses.

For data available at visit 9 (2 weeks post the third vaccination), positive response rates to Any Env will be compared between Groups 1 and 2 using the Barnard's exact test, where a positive response to Any Env is defined as those who are positive to at least one of the tested Env peptide pools. If there are more than 5 positive responders in both treatment groups, response magnitudes among positive responder will also be compared between Groups 1 and 2 using the Wilcoxon

rank sum test, where the response magnitude is defined as the summation across the Env PTE peptide pools. Statistical tests with a p-value ≤ 0.05 will be considered statistically significant.

10.5.4 Binding Antibody Multiplex Assay

IgG antibodies binding to antigens (gp 145 Ref. Std.; 1086C_D7gp120.avi/293F, Con S gp140 CFI, Con 6 g120/B, gp41, 1086C_V1_V2 Tags) will be measured using binding Ab multiplex assay (BAMA) as the primary analysis. Contingent on the primary analysis results of antigens listed above, a secondary set of IgG clone; H2 may be assessed. For the primary analysis, BAMA will be assayed for all samples at three time points (visits 2, 4 and 9). Specimens collected at visits 6 and/or 11 may also be assayed base on the results on the initial assay.

Serum HIV-1 specific IgG responses against these antigens are measured on 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, if provided. Samples from post-enrollment visits are declared to have positive responses if they meet three conditions: (1) based on the 95th percentile of the baseline visit serum samples and at least 100 MFI), (2) the MFI minus Neg values are greater than 3 times the baseline (visit 2) MFI minus Neg values, and (3) the MFI values are greater than 3 times the baseline MFI values. Background-adjusted MFI is used to summarize the magnitude at a given time-point. If sufficient immunogenicity is observed, samples may be titrated to calculate antibody titrations. Titrations are quantified by EC50/AUC.

The following analysis will be performed for each tested antigen separately. Positive response rates will be compared between treatment groups using the Barnard's exact test. Response magnitudes will be compared between treatment groups using the Wilcoxon rank sum test. Positive response rates will be compared within treatment group, across timepoints using the McNemar's test. Response magnitudes will be compared within treatment group, across timepoints using the Wilcoxon signed rank test. All p-values are two-sided. False-discovery-rate adjusted q-values [8] will be calculated to account for multiple antigens for each type of comparisons. Statistical significance is claimed if the unadjusted p-value is < 0.05 and the adjusted p-value (q-value) is < 0.2.

10.5.5 Neutralizing Antibody

Neutralizing antibodies against HIV-1 will be measured as a function of reductions in Tat-regulated luciferase (Luc) reporter gene expression in TZM-bl cells. Neutralization ID50 titers will be measured against Env vaccine strain (6980.v0.c22) and a single highly neutralization-sensitive Tier 1 virus as a positive control (MW965.26) for all available visit 9 samples. The global panel and/or clade-specific panels will be used to assess Tier 2 neutralization [6,7] initially in a subset of 15 participants randomly selected from Groups 1-3, and may expand to the remaining samples. Assays will be performed on specimens obtained at 2 weeks post 3th vaccination (visit 9) to assess the protocol secondary objective 3. Additional assays may be performed on the first day of vaccination (visit 2, baseline), 2 weeks post 1st vaccination (visit 4), 2 weeks post 2nd vaccination (visit 6), and 6 months post 3rd vaccination (visit 11, durability timepoint) for exploratory analyses.

A response is considered positive if the neutralization titer ID50 is above 10. . Magnitude of response will be measured by the natural log of the ID50 titer. An aggregate measure of response against the global panel will be calculated as the area-under-the-magnitude-breadth curve (AUC-MB), which is equivalent to the mean log titer across all viruses in the panel. If a titer is left censored, half the left censor limit will be used as the titer value.

Positive response rates are compared between treatment groups using the Barnard's exact test. Response magnitudes and AUC-MBs are compared between treatment groups using the Wilcoxon rank sum test. If data are available from multiple time-points, positive response rates are compared within treatment group, across timepoints using the McNemar's test andresponse magnitudes using the Wilcoxon signed rank test. Statistical tests with a p-value ≤ 0.05 will be considered statistically significant.

11 SAFETY TABLES, PARTICIPANT LISTINGS, AND FIGURES

11.1 List of Tables

The following tables are included in the SMB reports and FSR for Safety:

- Enrollment Report
- Demographics and Vaccination Frequencies
- Overall Protocol Status
- Maximum Local and Systemic Reactogenicity Summaries
- Adverse Experiences by Body System and Severity – By Decreasing Frequency
- Adverse Experiences by Preferred Term and Severity – By Decreasing Frequency – Includes Severe, Life-threatening or Fatal Experiences Only
- Adverse Experiences by Preferred Term and Severity – By Decreasing Frequency – Includes Experiences of All Severities
- Adverse Experiences by Preferred Term and Relationship to Study Product – By Decreasing Frequency – Includes Related Experiences Only
- Adverse Experiences by Preferred Term and Relationship to Study Product – By Decreasing Frequency – Includes Experiences of Any Relationship
- Expedited Adverse Experiences (EAEs) Reported to the Regulatory Support Center (RSC)
- Pregnancy Listing

11.2 List of Participant Listings

These participant listings are included in the SMB reports:

- Discontinuation Status
- Pregnancies
- Severe or Life-Threatening Local and Systemic Reactogenicities
- Moderate or Severe Erythema and Induration
- Expedited Adverse Experiences (EAEs)
- Severe, Life-Threatening, or Fatal Adverse Experiences
- Adverse Experiences with Relationship to Study Product
- HIV Infection Results from Lab and Reported by Site
- Study Product Administration Error

11.3 List of Figures

These graphs are included in the SMB reports and FSR for Safety:

- Maximum Local Reactogenicities
- Maximum Systemic Reactogenicities
- Boxplots for WBC, neutrophils, lymphocytes, hemoglobin, platelets, ALT, and creatinine

12 IMMUNOGENICITY TABLES AND FIGURES, BY ASSAY

12.1 Intracellular Cytokine Staining

Cellular immune responses will be measured by ICS using the following synthetic peptide pools: Env-1-PTEG-SEQ, Env-2-PTEG-SEQ, Env-3A-PTEG-SEQ, Env-3B-PTEG-SEQ. The following functional markers will be measured from a 17-colormetric panel: CD4 BUV39, CD3 BUV737, IFN-γ V450, AViD, CD14 BV510, CD56 BV570, PD-1 (CD279) BV605, CD8 BV650, ICOS (CD278) BV711, CCR7 BV786, TNF-α FITC, IL-4 PerCP-Cy5.5, IL-2 PE, CXCR5 PE-Dazzle594, IL-17a PE-Cy7, CD154 APC, Granzyme B Alx700, CD45RA APC-H7.

ICS will be performed on specimens from all participants collected at 2 weeks post 3rd vaccination (visit 9) to assess the protocol secondary objective 3. Additional assays may also be performed on specimens collected at 2 weeks post 1st vaccination (visit 4), at 2 weeks post 2nd vaccination (visit 6) at 6 months post 3rd vaccination (visit 11, durability/memory timepoint) for exploratory analyses.

12.1.1 List of Tables

- Response rate table by lab, T-cell subset, peptide pool, visit, day, and group/treatment arm for cells expressing either IL-2 and/or IFN-γ using Fisher's exact test criteria.
- Listing of positive responders for cells expressing either IL-2 and/or IFN-γ.
- Summary statistics (i.e., min, mean, median, max) among responders for T-cell subset, peptide pool, visit, and group/treatment arm

- Summary statistics (i.e., min, mean, median, max) among all participants (positive and negative responders) for T-cell subset, peptide pool, visit, and group/treatment arm.
- Response rate and/or response magnitude comparison of treatment arms.

12.1.2 List of Graphs

- Boxplots of background-adjusted IFN-γ and/or IL-2 response magnitude by T-cell subset, HIV protein, visit, and group/treatment arm using the Fisher exact test.
- Boxplots of background-adjusted marginal functional marker response magnitude by T-cell subsets, HIV genes, visits, and groups if requested. The default layout of boxplots will be with all available time points by treatment group on the same graph, one T-cell subset per graph.
- Boxplots of cytokine degree by T-cell subset, HIV protein, visit, and group/treatment arm if polyfunctional graphs are requested. The default layout of boxplots will be with all available time points by treatment group on the same graph, one T-cell subset per graph.
- Boxplots for each cytokine combination of a given degree by T-cell subset, HIV protein, visit, and group/treatment arm if polyfunctional graphs are requested. The default layout of boxplots will be with all available time points by treatment group on the same graph, one T-cell subset per graph.
- If assays are run for visit 4, visit 6 and/or 11, spaghetti plots will be produced to show IFN-γ and/or IL-2 response magnitude over time, with one graph per T-cell subset and HIV gene, with separate panels for treatment groups. The default layout of boxplots will be with all available time points by treatment group on the same graph, one T-cell subset per graph.

12.2 Binding Antibody Multiplex Assay

IgG antibodies binding to antigens (gp 145 Ref. Std.; 1086C_D7gp120.avi/293F, Con S gp140 CFI, Con 6 g120/B, gp41, 1086C_V1_V2 Tags) will be measured using binding Ab multiplex assay (BAMA) as the primary analysis. Contingent on the primary analysis results of antigens listed above, a secondary set of IgG clone; H2 may be assessed. For the primary analysis, BAMA will be assayed for all samples at three time points (visits 2, 4 and 9). Specimens collected at visits 6 and/or 11 may also be assayed base on the results on the initial assay.

12.2.1 List of Tables

- Response rate table by lab, isotype, antigen, visit, day, and group/treatment arm
- Listing of magnitudes (MFI*) among positive responders.
- Summary statistics (i.e., min, mean, median, max) among responders by lab, isotype, antigen, visit, day, and group/treatment arm.
- Summary statistics (i.e., min, mean, median, max) among all participants (positive and negative responders) by lab, isotype, antigen, visit, day, and group/treatment arm.
- Comparisons of response rates between treatment groups at each visit.
- Comparisons of response magnitudes among responders between treatment groups at each visit, if there are at least 5 responders within each group at a given visit.
- Comparisons of response rates between visits 4 and 9 for each treatment group.

- Comparisons of response magnitudes among responders between visits 4 and 9 for each treatment group if there are at least 5 responders at each visit within a given treatment group.

12.2.2 List of Graphs

- Barplots of response rates by isotype, antigen, visit day, and group/treatment arm.
- Spaghetti plot of background-adjusted MFI over time by antigen, antibody class (if applicable), visit, sample type and group, with boxplots among responders overlaid.

12.3 Neutralizing Antibody

Neutralization ID50 titers will be measured against Env vaccine strain (6980.v0.c22) and a single highly neutralization-sensitive Tier 1 virus as a positive control (MW965.26). The global panel of 9 viruses: 246F3, Ce1176, CNE55, X1632, Ce0217, BJOX2000, 25710, TRO11, CH119 are used to assess Tier 2 neutralization [6,7]. Assays will be performed on specimens obtained at 2 weeks post 3th vaccination (visit 9) to assess the protocol secondary objective 3. Additional assays may be performed on the first day of vaccination (visit 2, baseline), 2 weeks post 1st vaccination (visit 4), 2 weeks post 2nd vaccination (visit 6), and 6 months post 3rd vaccination (visit 11, durability timepoint) for exploratory analyses.

12.3.1 List of Tables

- Response rate table by lab, virus tier, virus, visit, and group/treatment arm.
- Listing of positive responders based on the criteria used for the response rate table(s).
- Summary statistics (i.e., min, geometric mean, median, max) of ID50 titers among responders by lab, virus tier, virus, visit, and group/treatment arm
- Summary statistics (i.e., min, geometric mean, median, max) of ID50 titers among all participants (positive and negative responders) by lab, virus tier, virus, visit, and group/treatment arm
- Comparisons of response rates between Group 1 vs. Group 3, Group 2 vs. Group 3, and Group 1 vs. Group 2 for each of the non-global panel viruses
- Comparisons of ID50 titers between Group 1 vs. Group 2 for each of the non-global panel viruses among responders.
- Comparisons of ID50 titers between Group 1 vs. Group 2 for each of the non-global panel viruses among all participants, with the titer of non-responders censored at 5.
- Comparison of AUC-MB between Group 1 vs. Group 2.

12.3.2 List of Graphs

- Boxplots of neutralizing antibody titers by virus, and group/treatment arm for the non-global panel viruses
- Boxplots of neutralizing antibody titers by virus and group/treatment arm for the global panel viruses
- If assays are run for visits 2, 4, 6 and/or 11, spaghetti plots of neutralizing antibody titer over time by cell type, isolate and group

- Magnitude-breadth (M-B) plots of titer and breadth for the global panel of 9 viruses by group/treatment arm.

13 REFERENCES

1. Agresti A, Coull BA. Approximate is better than "exact" for interval estimation of binomial proportions. Am Stat **1998**;52:119-26.
2. Lydersen S, Fagerland MW, Laake P. Recommended tests for association in 2 x 2 tables. Stat Med 2009;28:1159-75.
3. James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning with applications in R. New York: R. Springer, 2013.
4. Finak G, McDavid A, Chattopadhyay P, et al. Mixture models for single-cell assays with applications to vaccine studies. Biostatistics 2014;15:87-101.
5. Lin L, Finak G, Ushey K, et al. COMPASS identifies T-cell subsets correlated with clinical outcomes. Nat Biotechnol 2015;33:610-6.
6. Seaman MS, Janes H, Hawkins N, et al. Tiered categorization of a diverse panel of HIV-1 Env pseudoviruses for assessment of neutralizing antibodies. J Virol 2010;84:1439-52.
7. DeCamp A, Hraber P, Bailer RT, et al. Global panel of HIV-1 Env reference strains for standardized assessments of vaccine-elicited neutralizing antibodies. J Virol 2014;88:2489-507.
8. Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B 57, 289–300.