STATISTICAL ANALYSIS PLAN

A phase 1 clinical trial to evaluate the safety and immunogenicity of HIV-1 BG505 SOSIP.664 gp140 with TLR agonist and/or Alum adjuvants and VRC HIV Env Trimer 4571 and 3M-052-AF with Alum in healthy, HIV-uninfected adults

HVTN 137 (Protocol Version 5.0)

Date SAP finalized for signature: 11 February 2025

Effective Date: Date of last signature

SAP Version: 8.0

STATISTICAL ANALYSIS PLAN

Protocol Name:	A phase 1 clinical trial to evaluate the safety and immunogenicity of HIV-1 BG505 SOSIP.664 gp140 with TLR agonist and/or Alum adjuvants and VRC HIV Env Trimer 4571 and 3M-052-AF with Alum in healthy, HIV-uninfected adults
Protocol Number:	HVTN 137
SAP Author(s):	Lorel Schmitzberger, Statistical Research Associate III
SAP Version:	8.0

For acceptable methods for applying a signature, please refer to *SOP-0013 Statistical Analysis Plans section 5.4.1.1.*

Author(s):

Statistical Research Associate	
Legal Name	Lorel Schmitzberger
Job Title	Statistical Research Associate III
Signature & Date	See eTMF signature manifest

Approval(s):

Protocol Statistician	
Legal Name	Holly Janes
Job Title	Professor
Signature & Date	See eTMF signature manifest

The effective date of this document is the date of the latest signature.

TABLE OF CONTENTS

1. OVERVIEW

The following describes the Statistical Analysis Plan (SAP) for the analysis of data from HVTN 137 for Safety Monitoring Board (SMB) reports, the Final Study Report (FSR) for Safety, Protocol Team (PT) reports for immunogenicity data, and the FSR for Immunogenicity. As detailed in SCHARP SOP-0013, Revision 7 (effective date: October 1, 2019), this SAP is required prior to the first analysis and must be approved by the lead protocol statistician. SMB reporting begins shortly after enrollment opens, and subsequent revisions are expected to describe analysis of immunogenicity data. The SAP will be reviewed prior to the first SMB report and before the final analysis with all major revisions of the plan archived.

2. PROTOCOL SUMMARY

Title

A phase 1 clinical trial to evaluate the safety and immunogenicity of HIV-1 BG505 SOSIP.664 gp140 with TLR agonist and/or Alum adjuvants and VRC HIV Env Trimer 4571 and 3M-052-AF with Alum in healthy, HIV-uninfected adults.

Study products and routes of administration

- BG505 SOSIP.664 gp140: A stable, soluble cleaved trimeric HIV envelope (Env) gp140 glycoprotein based on the HIV-1 subtype A strain BG505 envelope sequence. To be administered as 100 mcg dose admixed with different adjuvants as 0.5 mL IM injection in the deltoid muscle.
- Trimer 4571 (VRC-HIVRG096-00-VP) (labeled as HIV-1 Trimer 4571 Vaccine): Consists of an HIV-1 envelope (Env) trimer variant derived from the clade A HIV-1 strain BG505. To be administered IM as 100-mcg dose admixed with 5 mcg of 3M-052-AF and 500 mcg of Alum.
- CpG 1018 adjuvant (labeled as CpG 1018 Drug Product): A cytosine phosphoguanosine oligodeoxynucleotide (CpG-ODN) designated 1018 containing a short unmethylated immunostimulatory DNA sequence (CpG), a toll-like receptor (TLR) 9 agonist. To be administered IM as 300 mcg dose admixed with BG505 SOSIP.664 gp140 and 500 mcg of Alum.
- 3M-052-AF adjuvant (labeled as AP 60-702): An aqueous formulation (AF) of the small molecule imidazoquinoline immune response modifier (IRM) 3M-052; TLR7/8 agonist. To be administered IM as 1 mcg or 5 mcg admixed with BG505 SOSIP.664 gp140 and 500 mcg of Alum in Part A and at the highest tolerated dose from Part A admixed with BG505 SOSIP.664 gp140 and 500 mcg of Alum in Part B. To be administered IM as 3 mcg with BG505 SOSIP.664 gp140 and 500 mcg of Alum (Group 7) or 5 mcg with Trimer 4571 and 500 mcg of Alum (Group 8) in Part C
- GLA-LSQ adjuvant (labeled as AP 10-602): A liposomal formulation of the synthetic TLR4 ligand glucopyranosyl lipid A (GLA) with the saponin *Quillaja saponaria* fraction 21 (QS-21). To be administered IM as 5 mcg GLA and 2 mcg QS-21 admixed with BG505 SOSIP.664 gp140.

- Alum adjuvant (labeled as Aluminum Hydroxide Suspension): Aluminum hydroxide to be administered IM as 500 mcg (aluminum content) admixed with the BG505 SOSIP.664 gp140 and Trimer 4571 (and adjuvants) as shown.
- Placebo: Tris-NaCl Diluent.

Schema

	N (V:P)	Protein Dose	Adjuvant dose	Injection schedule in months (days)		
				M0 (D0)	M2 (D56)	M6 (D168) Optional 2nd Boost
Part A 3M-052 dose escalation						
Group 1	5:1	100 mcg	1 mcg 3M-052-AF + 500 mcg Alum	BG505 SOSIP.664 gp140 + (3M-052-AF + Alum)	BG505 SOSIP.664 gp140 + (3M-052-AF + Alum)	BG505 SOSIP.664 gp140 + (3M-052-AF + Alum)
Group 2	10:1	100 mcg	5 mcg 3M-052-AF + 500 mcg Alum	BG505 SOSIP.664 gp140 + (3M-052-AF + Alum)	BG505 SOSIP.664 gp140 + (3M-052-AF + Alum)	BG505 SOSIP.664 gp140 + (3M-052-AF + Alum)
Part B Adjuvant comparison						
Group 3	20:2	100 mcg	300 mcg CpG 1018 + 500 mcg Alum	BG505 SOSIP.664 gp140 + (CpG 1018 + Alum)	BG505 SOSIP.664 gp140 + (CpG 1018 + Alum)	BG505 SOSIP.664 gp140 + (CpG 1018 + Alum)
Group 4	20:2	100 mcg	5 mcg 3M-052-AF + 500 mcg Alum	BG505 SOSIP.664 gp140 + (3M-052-AF + Alum)	BG505 SOSIP.664 gp140 + (3M-052-AF + Alum)	BG505 SOSIP.664 gp140 + (3M-052-AF + Alum)
Group 5	20:2	100 mcg	GLA-LSQ (5 mcg GLA + 2 mcg QS-21)	BG505 SOSIP.664 gp140 + GLA-LSQ	BG505 SOSIP.664 gp140 + GLA-LSQ	BG505 SOSIP.664 gp140 + GLA-LSQ
Group 6	20:2	100 mcg	500 mcg Alum	BG505 SOSIP.664 gp140 + Alum	BG505 SOSIP.664 gp140 + Alum	BG505 SOSIP.664 gp140 + Alum
Part C 3M052-AF + Alum Optimization for Elicitation of Neutralizing Activity with Trimers						
Group 7	10:1	100 mcg	3 mcg 3M-052-AF + 500 mcg Alum	BG505 SOSIP.664 gp140 + (3M-052-AF + Alum)	BG505 SOSIP.664 gp140 + (3M-052-AF + Alum)	BG505 SOSIP.664 gp140 + (3M-052-AF + Alum)
Group 8	10:1	100 mcg	5 mcg 3M-052-AF + 500 mcg Alum	Trimer 4571 + (3M-052-AF + Alum)	Trimer 4571 + (3M-052-AF + Alum)	Trimer 4571 + (3M-052-AF + Alum)
Total	**127: 115 vaccine recipients / 12 placebo recipients**					

3. OBJECTIVES AND ENDPOINTS

3.1 Objectives and endpoints for Part A

3.1.1 Primary objectives and endpoints for Part A

Primary objective 1:

To evaluate the safety and tolerability of intramuscular (IM) administration of BG505 SOSIP.664 gp140 with 3M-052-AF + Alum

Primary endpoint 1:

Signs and symptoms of local and systemic reactogenicity, laboratory measures of safety, adverse events (AEs), expedited AEs (EAEs), and AEs of special interest (AESIs)

3.1.2 Secondary objectives and endpoints for Part A

Secondary objective 1:

To evaluate and compare immune responses elicited by BG505 SOSIP.664 gp140 with each 3M-052-AF +Alum dose regimen

Secondary endpoints 1:

Response rate and magnitude of serum antibody neutralization of autologous (HIV-1 BG505.664) and other heterologous Tier 2 HIV-1 strains as measured by the TZM-bl assay 2 weeks after the second and after the optional third vaccination, if available

Response rate and magnitude of serum IgG binding antibodies to BG505 SOSIP.664 gp140 as measured by BAMA 2 weeks after the second and after the optional third vaccination, if available

Response rate and magnitude of IgG+ B cells binding to BG505 SOSIP.664 gp140 tetramers as measured by multiparameter flow cytometry 2 weeks after the second and after the optional third vaccination, if available

Percent HIV Env-specific CD4+ T cells in blood as measured by multiparameter flow cytometry 2 weeks after the second and after the optional third vaccination, if available

3.1.3 Exploratory objectives for Part A

Exploratory objective 1:

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

Exploratory objective 2:

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

3.2 Objectives and endpoints for Part B

3.2.1 Primary objectives and endpoints for Part B

Primary objective 1:

To evaluate the safety and tolerability of intramuscular (IM) administration of BG505 SOSIP.664 gp140 with the following adjuvants: CpG 1018 + Alum, 3M-052-AF + Alum, GLA-LSQ, or Alum alone

Primary endpoint 1:

Signs and symptoms of local and systemic reactogenicity, laboratory measures of safety, adverse events (AEs), expedited AEs, and AEs of special interest (AESIs)

Primary objective 2:

To evaluate and compare the neutralizing antibody response induced by BG505 SOSIP.664 gp140 with different adjuvants (CpG 1018 + Alum, 3M-052-AF + Alum, GLA-LSQ, or Alum alone)

Primary endpoint 2:

Response rate and magnitude of serum antibody neutralization of autologous (HIV-1 BG505) and other heterologous Tier 2 HIV-1 strains as measured by TZM-bl assay 2 weeks after the second and third vaccinations

3.2.2 Secondary objectives and endpoints for Part B

Secondary objective 1:

To evaluate and compare the durability of neutralizing antibody responses elicited by BG505 SOSIP.664 gp140 with various adjuvants (CpG 1018 + Alum, 3M-052-AF + Alum, GLA-LSQ, or Alum)

Secondary endpoint 1:

Response rate and magnitude of serum antibody neutralization of autologous (HIV-1 BG505) and other heterologous Tier 2 HIV-1 strains as measured by TZM-bl assay 6 and 12 months post third vaccination

Secondary objective 2:

To evaluate and compare the peak and durability of serum binding antibody responses elicited by BG505 SOSIP.664 gp140 HIV-1 Env protein with various adjuvants (CpG 1018 + Alum, 3M-052-AF + Alum, GLA-LSQ, or Alum)

Secondary endpoints 2:

Response rate, magnitude, and breadth of IgG binding antibody responses to HIV BG505 SOSIP.664 gp140, and HIV-1 gp120 and gp140 variant strains (to determine off-target non-neutralizing Abs) as measured by BAMA 2 weeks after the second vaccination and 2 weeks, 6 months, and 12 months after the third vaccination

Secondary objective 3:

To evaluate and compare the peak and durability of BG505 SOSIP.664 gp140-specific B cells elicited by BG505 SOSIP.664 gp140 with various adjuvants (CpG 1018 + Alum, 3M-052-AF + Alum, GLA-LSQ, or Alum)

Secondary endpoint 3:

Response rate and magnitude of memory IgG+ B cells binding to BG505 SOSIP.664 gp140 tetramers as measured by multiparameter flow cytometry 2 weeks after the second vaccination and 2 weeks, 6 months, and 12 months after the third vaccination

Response rate and magnitude of BG505 SOSIP.664 gp140-specific plasmablasts as measured by multiparameter flow cytometry 1 week after the second and third vaccinations

Secondary objective 4:

To evaluate Env-specific CD4+ T-helper subpopulations, phenotypes and functions elicited by BG505 SOSIP.664 gp140 with various adjuvants (CpG 1018 + Alum, 3M-052-AF + Alum, GLA-LSQ, or Alum)

Secondary endpoint 4:

Percent HIV Env-specific cytokine-expressing CD4+ T cells in blood by intracellular cytokine staining (ICS) multiparameter flow cytometry and polyfunctional subset analysis 2 weeks after the second vaccination and 2 weeks, 6 months, and 12 months after the third vaccination

Secondary objective 5:

To evaluate innate immune responses and baseline predictors, compare the innate immune responses elicited by BG505 SOSIP.664 gp140 with various adjuvants (CpG 1018 + Alum, 3M-052-AF + Alum, GLA-LSQ, or Alum) after the first vaccination, and correlate with adaptive immune responses

Secondary endpoints 5:

Alterations in blood leukocyte populations during the innate response (day 1, 3 and 7) relative to prevaccine levels (day 0)

Alterations in RNAseq expression of leukocyte and/or immune cells (lymphocyte populations, natural killer (NK) cells, dendritic cell (DC) subsets, monocytes subsets, and granulocytes) (days 1, 3, and 7) relative to prevaccine levels (day 0)

Alterations of concentrations of immune cytokines and chemokines in serum samples postvaccination (days 1, 3, 7) relative to prevaccine levels (day 0)

Secondary objective 6:

To characterize systemic inflammatory markers among participants with moderate to severe reactogenicity after any vaccination

Secondary endpoint 6:

Blood cell subpopulation dynamics by multiparameter flow cytometry, gene expression alterations by RNAseq/transcripts and soluble cytokine/inflammatory mediator alterations comparing day of last vaccination and day of visit for moderate to severe reactogenicity

3.2.3 Exploratory objectives for Part B

Exploratory objective 1:

To further characterize antibody responses (eg, IgA/IgG subclass binding, epitope mapping, Ab avidity, infected cell binding, linear peptide array)

Exploratory objective 2:

To characterize Fc-mediated antibody functions (eg, fragment crystallizable receptor [FcR] array, antibody-dependent cellular phagocytosis [ADCP], antibody-dependent neutrophil phagocytosis [ADNP], antibody-dependent cellular cytotoxicity [ADCC], and/or complement binding/deposition)

Exploratory objective 3:

To further characterize neutralizing antibody responses using diagnostic mutants of HIV-1 Env pseudotyped viruses in the TZM-bl assay

Exploratory objective 4:

To analyze the B cell repertoire including somatic hypermutation, affinity maturation, CDRH3 length, and other signatures of bnAb class epitopes

Exploratory objective 5:

To apply systems biology approaches (eg, RNAseq, LC-MS) to examine transcriptomic, and metabolomic and/or proteomic profiles of innate and adaptive immune responses

Exploratory objective 6:

To characterize Env-specific B and CD4+ T cells (including Tfh/pTfh) and antigen receptor sequences using samples such as PBMC, lymph node fine needle aspirates, bone marrow aspirates, or leukapheresis

Exploratory objective 7:

To assess polyclonal serum Ab (Fab) binding to Env epitopes using electron microscopic imaging

Exploratory objective 8:

To measure Env-specific antibody levels in mucosal secretions and tissues using a quantitative immunoassay and/or immunohistochemistry (IHC)

Exploratory objective 9:

To identify the regions and epitopes of the BG505 SOSIP.664 gp140 protein targeted by CD4+ T cells in response to vaccination

Exploratory objective 10:

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

Exploratory objective 11:

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

3.3 Objectives and endpoints for Part C

3.3.1 Primary objectives and endpoints for Part C

Primary objective 1:

To evaluate the safety and tolerability of IM administration of Trimer 4571 (BG505 SOSIP.664 DS gp140) with 3M-052-AF + Alum.

Primary endpoint 1:

Signs and symptoms of local and systemic reactogenicity and laboratory measures of safety, AEs, expedited AEs, and AESIs.

Primary objective 2:

To evaluate the neutralizing antibody response induced by Trimer 4571 (BG505 SOSIP.664 DS gp140) with 3M-052-AF + Alum.

Primary endpoint 2:

Response rate and magnitude of serum antibody neutralization of autologous (HIV-1 BG505) and other heterologous tier-2 HIV-1 strains, as measured by TZM-bl assay 2 weeks after the second and third vaccinations.

3.3.2 Secondary objectives and endpoints for Part C

Secondary objective 1:

To test whether a 3-mcg dose of 3M-052-AF + Alum is sufficient to generate autologous serum antibody neutralization (HIV-1 BG505) and other heterologous tier-2 HIV-1 strains, as measured by TZM-bl assay 2 weeks after the second and third vaccinations.

Secondary endpoint 1:

Response rate and magnitude of serum antibody neutralization of autologous (HIV-1 BG505) and other heterologous tier-2 HIV-1 strains, as measured by TZM-bl assay 2 weeks after the second and third vaccinations, comparing the 5-mcg dose of the 3M-052-AF + Alum with the 3-mcg dose of 3M-052-AF + Alum.

Secondary objective 2:

To evaluate and compare the peak and durability of serum binding antibody responses elicited by BG505 SOSIP.664 gp140 adjuvanted with 3 mcg of 3M-052-AF+Alum, and by Trimer 4571 adjuvanted with 5 mcg of 3M-052-AF+Alum with Part B groups

Secondary endpoints 2:

Response rate, magnitude, and breadth of IgG binding antibody responses to HIV BG505 SOSIP.664 gp140, Trimer 4571, and HIV-1 gp120 and gp140 variant strains (to determine off-target non-neutralizing Abs) as measured by BAMA 2 weeks after the second vaccination and 2 weeks, 6 months, and 12 months after the third vaccination

Secondary objective 3:

To evaluate and compare the peak and durability of BG505 SOSIP.664 gp140-specific B cells elicited by BG505 SOSIP.664 gp140 adjuvanted with 3 mcg of 3M-052-AF+Alum, and by Trimer 4571 adjuvanted with 5 mcg of 3M-052-AF+Alum with Part B groups

Secondary endpoint 3:

Response rate and magnitude of memory IgG+ B cells binding to BG505 SOSIP.664 gp140 tetramers as measured by multiparameter flow cytometry 2 weeks after the second vaccination and 2 weeks, 6 months, and 12 months after the third vaccination

Response rate and magnitude of BG505 SOSIP.664 gp140 antigen-specific plasmablasts as measured by multiparameter flow cytometry 1 week after the second and third vaccinations

Secondary objective 4:

To evaluate Env-specific CD4+ T-helper subpopulations, phenotypes and functions elicited by BG505 SOSIP.664 gp140 adjuvanted with 3 mcg of 3M-052-AF+Alum, and by Trimer 4571 adjuvanted with 5 mcg of 3M-052-AF+Alum

Secondary endpoint 4:

Percent HIV Env-specific cytokine-expressing CD4+ T cells in blood by intracellular cytokine staining (ICS) multiparameter flow cytometry and polyfunctional subset analysis 2 weeks after the second vaccination and 2 weeks, 6 months, and 12 months after the third vaccination

Secondary objective 5:

To evaluate innate immune responses and baseline predictors, compare the innate immune responses elicited by BG505 SOSIP.664 gp140 adjuvanted with 3 mcg of 3M-052-AF+Alum, and by Trimer 4571 adjuvanted with 5 mcg of 3M-052-AF+Alum after the first vaccination, and correlate with adaptive immune responses

Secondary endpoints 5:

Alterations in blood leukocyte populations during the innate response (day 1, 3 and 7) relative to prevaccine levels (day 0)

Alterations in RNAseq expression of leukocyte and/or immune cells (lymphocyte populations, natural killer (NK) cells, dendritic cell (DC) subsets, monocytes subsets, and granulocytes) (days 1, 3, and 7) relative to prevaccine levels (day 0)

Alterations of concentrations of immune cytokines and chemokines in serum samples postvaccination (days 1, 3, 7) relative to prevaccine levels (day 0)

Secondary objective 6:

To characterize systemic inflammatory markers among participants with moderate to severe reactogenicity after any vaccination

Secondary endpoint 6

Blood cell subpopulation dynamics by multiparameter flow cytometry, gene expression alterations by RNAseq/transcripts and soluble cytokine/inflammatory mediator alterations comparing day of last vaccination and day of visit for moderate to severe reactogenicity

3.3.3 Exploratory objectives and endpoints for Part C

Exploratory objective 1:

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

Exploratory objective 2:

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

4. COHORT DEFINITION

Participants

Recruitment will target enrolling 127 healthy, HIV-1–uninfected volunteers aged 18 through 50 years (inclusive); 115 vaccine recipients and 12 placebo recipients.

Design

Multicenter, randomized, controlled, double-blind trial

Duration per participant

Part A: 8 months of scheduled clinic visits (main study) plus an adverse event of special interest (AESI) health contact at month 14

Part A Optional Second Boost: additional 12 months of scheduled clinic visits (main study) after the third vaccination

Part B: 18 months of scheduled clinic visits (main study)

Part C: 18 months of scheduled clinic visits (main study)

Estimated total study duration

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

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

A participant's randomization assignment will be computer generated and provided to the HVTN CRS pharmacist through a Web-based randomization system. For Part A, B and C, participants within each dose group will be randomized to vaccine or placebo assignment. For Part C, Group 7 will be enrolled before Group 8. There will be no randomization for Groups 1 and 2 as they will be enrolled sequentially for dose escalation. 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]).

7. BLINDING

Participants and site staff (except for site pharmacists) will be blinded as to participant group and to treatment arm assignments and, in Part B to participant group assignments as well. 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.

In some cases, the CRS, PSRT, or study sponsor may believe unblinding of the site Principal Investigator (PI) and participant would be appropriate to facilitate the clinical management of an AE or SAE. The HVTN Unblinding MOP specifies procedures for emergency unblinding, and for early unblinding for medical reasons.

8. SAMPLE SIZE AND POWER

8.1 Sample size calculations for safety (Part A)

The purpose of Part A is to select the 3M-052-AF adjuvant dose, by identifying potential 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 the low-dose vaccine arm (n = 5), there is a 90% chance of observing at least 1 event if the true rate of such an event is 37% or more; and there is at least a 90% chance of observing no events if the true rate is 1% or less. For the high-dose vaccine arm (n = 10), there is at least a 90% chance of observing at least 1 event if the true rate of such an event is 21% or more; and there is a 90% chance of observing no events if the true rate is 1% or less. For the two vaccine arms combined (n = 15), there is at least a 90% chance of observing at least 1 event if the true rate of such an event is 14.3% or more; there is a 90% chance of observing no events if the true rate is 0.5% or less. As a reference, in HVTN vaccine trials from April 2008 through March 2018, about 1.74% of participants who received placebos experienced an SAE.

Binomial probabilities of observing 0, 1 or more, and 2 or more events among arms of size 5 and 10 are presented in Table 1 for a range of possible true adverse event rates. These calculations provide a more complete picture of the sensitivity of the Part A study design to identify potential safety problems with the 3M-052-AF-adjuvanted vaccine.

Table 1: Probability of observing 0 event, 1 or more events, and 2 or more events, among arms of size 5 and 10, for different true event rates

True event rate (%)	Pr(0/5)	Pr(1+/5)	Pr(2+/5)	Pr(0/10)	Pr(1+/10)	Pr(2+/10)
1	95.1	4.9	0.1	90.4	9.6	0.4
4	81.5	18.5	1.5	66.5	33.5	5.8
10	59	41	8.1	34.9	65.1	26.4
20	32.8	67.2	26.3	10.7	89.3	62.4
30	16.8	83.2	47.2	2.8	97.2	85.1
40	7.8	92.2	66.3	0.6	99.4	95.4

8.2 Sample size calculations for safety (Part B)

The goal of the safety evaluation for Part B of this study is to identify safety concerns associated with administration of one or more vaccine regimens. Specifically, for each vaccine arm of Part B (n = 20), there is a 90% chance of observing at least 1 event if the true rate of such an event is 11% or more; and there is a 90% chance of observing no events if the true rate is 0.5% or less. For the four vaccine arms combined (n = 80), there is a 90% chance of observing at least 1 event if the true rate of such an event is 2.9% or more; and there is a 90% chance of observing no events if the true rate is 0.1% or less.

Binomial probabilities of observing 0, 1 or more, and 2 or more events among arms of size 20 and 80 are presented in Table 2 for a range of possible true adverse event rates. These calculations provide a more complete picture of the sensitivity of the Part B study design to identify potential safety problems with the four vaccine regimens.

Table 2: Probability of observing 0 event, 1 or more events, and 2 or more events, among arms of size 20 and 80, for different true event rates

True event rate (%)	Pr(0/20)	Pr(1+/20)	Pr(2+/20)	Pr(0/80)	Pr(1+/80)	Pr(2+/80)
1	81.8	18.2	1.7	44.8	5.2	19.1
4	44.2	55.8	19.0	3.8	96.2	83.5
10	12.2	87.8	60.8	0	100	99.8
20	1.2	98.8	93.1	0	100	100
30	0.1	99.9	99.2	0	100	100

An alternative way of describing the statistical properties of the study design is in terms of the 95% confidence interval for the true rate of an adverse event based on the observed data. Table 3 shows the 2-sided 95% confidence intervals for the probability of an event based on a particular observed rate. Calculations are done using the score test method [1]. If none of the 80 participants receiving a vaccine regimen experience a safety event, the 95% 2-sided upper confidence bound for the true rate of such events in the total vaccinated population is 4.6%. For each individual vaccine arm (n = 20), the 2-sided upper confidence bound for this rate is 16.1%.

Table 3: Two-sided 95% confidence intervals based on observing a particular rate of safety endpoints for arms of size n1 and n2

Observed event rate	95% confidence interval (%)
0/20	(0%, 16.1%)
1/20	(0.9%, 23.6%)
2/20	(2.8%, 30.1%)
0/80	(0%, 4.6%)
1/80	(0.2%, 6.7%)
2/80	(0.7%, 8.7%)

8.3 Sample size calculations for immunogenicity (Part B)

Immunogenicity sample size calculations allow for a 10% rate of missing immunogenicity data at Months 2.5 and 6.5 (n = 18 for analysis) and 15% rate of missing immunogenicity data at Months 12 and 18 (n = 17 for analysis).

The first goal is to evaluate immune response rates based on data from the primary TZM-bl neutralizing antibody assay among vaccinees. The precision with which the true response rate can be estimated from the observed data depends on the true underlying response rate and the sample size. Two-sided 95% confidence intervals for the response rate based on observing a particular response rate among vaccinees is shown in Table 4. Calculations are done using the score test method [1].

Table 4: Two-sided 95% confidence intervals for the true response rate based on observing a particular response rate among vaccinees (n = 18)

No. of responses	Observed response rate (%)	95% confidence interval
4/18	22.2	(9.0, 45.2)
6/18	33.3	(16.3, 56.3)

8/18	44.4	(24.6, 66.3)
10/18	55.6	(33.7, 75.4)
12/18	66.7	(43.7, 83.7)
14/18	77.8	(54.8, 91.0)
15/18	83.3	(60.8, 94.2)
16/18	88.9	(67.2, 96.9)
17/18	94.4	(74.2, 99.0)

As shown in Table 5, there is limited power for a formal comparison of immunogenicity response rates between vaccine arms of size n = 20. For either 80% or 90% power, the sizes of differences that the trial is powered to detect are large. These calculations use a Fisher's exact 2-sided test with a Type I error rate of 0.05. However, the study is adequately powered to detect the differences seen in the CAVD 641 study in rhesus macaques by Bali Pulendran. Rates of autologous tier 2 neutralizing antibody responses at week 26 (the peak timepoint) were compared between arms with Alum-3M-052 or GLA-LSQ adjuvant, both with BG505-SOSIP protein. The large difference in response rates – 50% vs. 100% – could be detected with 97% power.

Table 5: Power for comparison of response rates between 2 arms (n1 = 18, n2 = 18)

True response rate Arm 1 (%)	Minimum true response rate in Arm 2 in order to detect a difference	
	80% power	90% power
10	58	65
20	71	77
30	81	87
40	89	94
50	95	99

Power for comparing the magnitude of immune responses between pairs of arms with different adjuvants is first calculated using binding antibody data from the HEPLISAV study. Engerix, an HBV vaccine formulated with alum, was compared to HEPLISAV, the same HBV vaccine formulated with CpG 1018. With n = 18 vaccinees per arm, the current study has 100% power to detect the difference in geometric mean immune response seen in the HEPLISAV study at Week 28 (the peak timepoint) (79.7 mIU/mL vs. 206.1 mIU/mL), assuming based on the HEPLISAV study that the group with higher geometric mean has a smaller standard deviation (51.38 mIU/mL vs. 6.04 mIU/mL), using a 0.025-level one-sided Wilcoxon rank sum test. In addition, power is also 100% for detecting the difference seen between adjuvants at Week 52 (the durability timepoint), based on the HEPLISAV data (geometric means 19.0 vs. 131.0 mIU/mL and standard deviations 34.57 vs. 6.92 mIU/mL).

In addition, "generic" power calculations inform the power of the study to compare immune response magnitudes between arms. These presume a continuous immune response, transformed to a 1-standard deviation scale, and a mean shift in scaled immune response between arms. Table 6 shows the power for comparing immune responses between pairs of arms with different adjuvants, separately at Months 2.5 and 6.5 (n = 18 per arm) and Months 12 or 18 (durability, n = 17 per arm) time points. Observe that the study is powered to detect moderate 1-standard deviation differences in immune responses between adjuvants at the early timepoints and 1.1-standard-deviation differences at the durability timepoints.

Table 6: Power for comparing the magnitude of a generic immune response between arms. Immune responses are compared using 0.025-level one-sided Wilcoxon rank sum tests. Continuous immune responses are assumed to follow a normal distribution, transformed to a 1-standard deviation scale, with a mean of zero in one arm and the same standard deviation in the two arms. To allow for missing immunogenicity data, analyses include 17 or 18 subjects in each arm. Power is based on 1000 simulations.

Mean difference between arms (in standard-deviation units)	Power, n = 18 per arm	Power, n = 17 per arm
0.8 SD	63%	61%
0.9 SD	71%	69%
1.0 SD	81%	78%
1.1 SD	85%	85%
1.2 SD	92%	90%
1.3 SD	97%	95%

In addition, the ability of the study to rank and select the adjuvant with the highest immune response at a given timepoint was assessed. Table 7 shows the probability of correctly selecting the adjuvant with the highest mean immune response at Month 6.5, as a function of the mean in the arm with next-lowest mean immune response. Calculations again assume that immune responses have been transformed to the 1-standard deviation scale. The results show that the study is sufficiently sized to correctly rank adjuvant arms, assuming the top two arms have a mean difference of at least 0.5 standard-deviations.

Table 7: Probability of correct ranking. Probability of correctly selecting the adjuvant with the highest mean immune response at Month 6.5, based on a generic standard-deviation-scaled continuous immune response. Probability is shown as a function of the mean immune response in the next-best arm. To allow for missing immunogenicity data, analyses include 18 subjects in each arm. Calculations assume all arms below the best have a mean immune response equal to that of the next-best arm and thus the results are conservative.

Mean difference, best – next-best	Probability of correct ranking
0.1 SD	0.38
0.2 SD	0.52
0.3 SD	0.65
0.4 SD	0.75
0.5 SD	0.82
0.6 SD	0.92

8.4 Sample size calculations for safety (Part C)

The goal of the safety evaluation for Part C of this study is to identify safety concerns associated with administration of 1 or more vaccine regimens studied in Part C. Specifically, for each vaccine arm of Part C (n = 10), there is a 90% chance of observing at least 1 event if the true rate of such an event is 20.6% or more and there is a 90% chance of observing no events if the true rate is 1% or less.

Binomial probabilities of observing 0 events, 1 or more events, and 2 or more events among arms of size 10 are presented in Table 8 for a range of possible true adverse event rates. These calculations provide a more complete picture of the sensitivity of the Part C study design to identify potential safety problems with the two vaccine regimens.

Table 8: Probability of observing 0 event, 1 or more events, and 2 or more events, among arms of size 10, for different true event rates

True event rate (%)	Pr(0/10)	Pr(1+/10)	Pr(2+/10)
1	90.4	9.6	0.4
4	66.5	33.5	5.8
10	34.9	65.1	26.4
20	10.7	89.3	62.4
30	2.8	97.2	85.1

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 adverse event based on the observed data. Table 9 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 [1]. If none of the 10 participants receiving a vaccine regimen experience a safety event, the 95% 2-sided upper confidence bound for the true rate of such events in the total vaccinated population is 27.8%.

Table 9: Two-sided 95% confidence intervals based on observing a particular rate of safety endpoints for arms of size n1 and n2

Observed event rate	95% confidence interval (%)
0/10	(0%, 27.8%)
1/10	(1.8%, 40.4%)
2/10	(5.7%, 51%)

8.5 Sample size calculations for immunogenicity (Part C)

Immunogenicity sample-size calculations allow for a 10% rate of missing immunogenicity data at Month 2.5 and 6.5 (n = 9 for analysis).

The first goal is to evaluate immune response rates based on data from the primary TZM-bl neutralizing antibody assay among vaccinees. The precision with which the true response rate can be estimated from the observed data depends on the true underlying response rate and the sample size. Two-sided 95% CIs for the response rate based on observing a particular rate of responses in the vaccinees is shown in Table 10. Calculations are done using the score test method [1].

Table 10: Two-sided 95% confidence intervals for the true response rate based on observing a particular response rate among vaccinees (n = 9)

No. of responses	Observed response rate (%)	95% confidence interval
1/9	11.1	(2.0, 43.5)
2/9	22.2	(6.3, 54.7)
3/9	33.3	(12.1, 64.6)
4/9	44.4	(18.9, 73.3)
5/9	55.6	(26.7, 81.1)
6/9	66.7	(35.4, 87.9)
7/9	77.8	(45.3, 93.7)
8/9	88.9	(56.5, 98.0)
9/9	100.0	(70.1, 100.0)

As shown in Table 11and Table 12, there is limited power for a formal comparison of immunogenicity response rates between Parts B and C based on vaccine arms of size n = 10

(Parts A and C) and n = 20 (Part B). For either 80% or 90% power, the sizes of differences that the trial is powered to detect are large.

Table 11 Power for comparison of response rates between 2 arms (n1 = 9, n2 = 9)

True response rate Arm 1 (%)	Minimum true response rate in Arm 2 in order to detect a difference	
	80% power	90% power
10	83	91
20	92	99
30	99	N/A

Table 12 Power for comparison of response rates between 2 arms (n1 = 9, n2 = 18)

True response rate Arm 1 (%)	Minimum true response rate in Arm 2 in order to detect a difference	
	80% power	90% power
10	72	90
20	83	90
30	91	96
40	92	100

9. 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.

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, when the assay endpoint is viewed as a collection of hypotheses (e.g., testing multiple peptide pools to determine a positive response). Analyses will be performed using SAS and/or R.

9.1 Analysis variables

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

Another key variable has to do with implementation of the study. Specifically, given the conduct during the COVID-19 pandemic, study visit windows were extended to optimize visit attendance in Part A. Therefore “primary interval”, defined as days between first and second vaccinations, and “boost interval”, defined as days between second vaccination and boost, are key covariates in Part A.

9.2 Baseline comparability

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

9.3 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.

9.3.1 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 test will be used to test for differences in severity between arms.

9.3.2 AEs and SAEs

AEs will be summarized using Medical Dictionary for Regulatory Activities (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 the magnitude 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.

9.3.3 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 timepoint, as well as for changes between baseline and post-enrollment. 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 postvaccination timepoint. Reportable clinical laboratory abnormalities without an associated clinical diagnosis will also be included in the tabulation of AEs described above.

9.3.4 Reason 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.

9.4 Immunogenicity analysis

9.4.1 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 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.

Data from placebo arms will be pooled across groups (n = 8 for analysis in Part B). For analyses of immunogenicity after the second vaccination, data on the 3M-052-AF from Part A (with the maximum tolerated dose of the 3M-052-AF) and Part B will be pooled. Otherwise, analyses of immunogenicity will be performed separately for Part A and Part B participants.

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 (FPRs) in the control arms. Barnard's test 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 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 (CDFs) 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 pooled placebo group.

The difference between arms at a specific timepoint 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. Individual tests comparing the 6 pairs of vaccine arms will be performed unless prespecified. If rank-based tests are used, then the tests will be inverted to construct Hodges-Lehmann point estimates and 2-sided (1-0.05/6) × 100% confidence intervals (CIs) about the differences in location centers of the 6 pair-wise comparisons of vaccine arms. If actual-value tests are used then the Dunnett's procedure will be used to construct simultaneous confidence intervals about the pairs of mean differences for the many-to-one comparisons [2] when multiple vaccine arms are each compared with the pooled placebo group. Given that all pair-wise comparisons between the multiple vaccine arms are of interest, the Tukey procedure [3] will be used. An appropriate data transformation (e.g., $\log_{10}$ transformation) may be applied to better

satisfy assumptions of symmetry and homoscedasticity (constant variance). Significance of the differences between pairs will be evaluated using two procedures, first based on whether the simultaneous 95% CIs exclude zero and secondly based on whether the nominal (unadjusted) 95% CIs exclude zero.

Scatterplots will be used to visualize the association between quantitative immune responses at Month 2.5 and prime interval, and between quantitative immune responses at Month 6.5 and both prime and boost interval variables in Part A.

Some immunologic assays have underlying continuous or count-type readout that are dichotomized into responder/nonresponder categories (e.g., intracellular cytokine staining). If treatment arm differences for these assays are best summarized by a mixture model, then either Lachenbruch's test statistic [4] or an alternative two-part test [5] will be used to evaluate the composite null hypothesis of equal response rates in the 2 arms and equal response distributions. Lachenbruch's 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 (for example, linear mixed models or marginal mean models fit by generalized estimating equations) will be utilized to incorporate immune responses over several timepoints and to test for differences over time and differences in trajectories across treatment arms. For example, repeated measures analyses will be conducted to evaluate trajectories in innate immune response over Days 1, 3, 7 relative to responses measured at Day 0. 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 10% 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 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 are 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 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 evaluating repeated immunogenicity measurements over time, linear mixed effects model will be used. If the immunological outcomes are left- and/or right-censored, then the linear mixed effects models of Hughes [6] will be used, because they accommodate the censoring. In addition, secondary analyses of repeated immunogenicity measurements may be done using weighted generalized estimating equation (GEE) [7] 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.

Some immunogenicity endpoints are only measured in subset of participants. For example, fine needle lymph node aspirates, bone marrow aspirates, and leukapheresis will be performed on consenting participants at a subset of sites, and immunogenicity endpoints based on these samples will only exist for participants enrolled at those sites. For such endpoints, exploratory analyses will be conducted to assess the correlation of participant characteristics measured in (nearly) all participants with the resource-intensive endpoints. For example, if the same assay is performed on blood and fine needle lymph node aspirate samples, then a scatterplot and Spearman rank correlation coefficient (r) will be used to assess the correlation of responses. If at least moderate correlations exist (eg, $r \geq 0.3$), then the semiparametric efficient analysis method of Rotnitzky and Robins [8] will be used (described in Gilbert, Sato et al. for application to vaccine studies [9]) to estimate the mean of the resource-intensive endpoint for each group and to compare means between groups.

9.4.2 Multivariate display of immunogenicity endpoints

Data visualization techniques may be used to explore the relationship among immunogenicity readouts across assays, and across readouts for assays that produce high dimensional data. The set of readouts may be based on one of the primary endpoints (e.g., neutralizing Abs), 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 heatmap showing the degree of correlation between any two 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 [10]. 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 (e.g., Env-specific CD4+ T cell polyfunctionality score) from the origin to the point defined by its first two 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 two arrows conveys information about the correlation of the two readouts, with a zero 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.

9.4.3 Primary analyses of neutralization data

The primary measure of antibody neutralization is the response to the response to the vaccine-matched BG505.T332N isolate. As well, responses to a single highly neutralization-sensitive Tier 1 virus (MW965.26) will be assessed. The mean magnitude of the response to the vaccine-matched isolate will be estimated for each arm and compared to the pooled placebo group. As well, the ratio of the response to the vaccine-matched antigen, relative to that to the neutralization-sensitive isolate, will be estimated for each arm.

If a Tier 2 panel of viral isolates is used to assess the breadth of neutralizing antibody responses, the area-under-the-magnitude-breadth curve (AUC-MB) to a global panel of viral isolates [11] will be computed for each participant with evaluable neutralization data, as described in [12]. 95% CIs about the four differences in mean AUC-MB for each vaccine regimen versus the pooled placebo groups (vaccine – placebo), will be calculated.

9.4.4 Analysis of CD4+ T cell responses as measured by the ICS assay

The analysis of CD4+ 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 [13] 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. When 3 or more cytokines are being measured by the ICS assay, the polyfunctionality of ICS responses may also be analyzed as an exploratory endpoint. Besides descriptive plots of the magnitude of polyfunctional responses, the COMPASS (Combinatorial Polyfunctionality analysis of Antigen-Specific T-cell Subsets) statistical framework [14] 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.

9.4.5 Analysis of epitope mapping data

The minimal set of optimal length epitopes that can explain the observed T cells responses to a set of individually tested overlapping 15-mer peptides may be obtained to assess the breadth and depth of the T cell responses based on the epitope mapping data.

When human leukocyte antigen (HLA) data are also available, computational HLA:peptide binding predictors (e.g., NetMHCpan) may be used to more accurately identify each participant's T cell epitopes.

The set of epitopes may also be used to assess "coverage" of a representative set of circulating viruses by each participant's set of T-cell responses. The coverage provided by each epitope is defined by the fraction of viruses that match the sequence of the mapped epitope. The coverage provided by a participant's set of epitopes is defined by the fraction of viruses covered by at least one of the participant's epitopes.

9.4.6 Analysis of multiplexed immunoassay data

When a small panel of analytes (e.g. ≤ 5) is being assessed in a multiplexed immunoassay, the analysis of response rates and response magnitudes will be evaluated and compared under the general approach. When a larger panel is being assessed, 3 approaches may be considered to evaluate the magnitude and breadth of these responses. First, Magnitude–Breadth (M-B) curves maybe employed to display individual- and group-level response breadth as a function of magnitude. Applied to Tier 2 virus neutralizing antibody data, response breadth is defined as the number of isolates with magnitude above a positivity threshold. Two choices are to compare the M-B curves among vaccine arms, as follows: a non-parametric Wilcoxon rank sum test on the subject-specific AUC-MB or a Kolmogorov-Smirnov type test on the 2 group-average M-B curves. Simulations can be used to obtain 2-sided p-values for the latter test. Second, a weighted-average score-like variable may be constructed to account for the correlations between analytes as an integrate magnitude of responses to multiple analytes. Similar group comparison methods described in the first approach may be adopted. Third, machine-learning approaches such as super-learning [15] will be employed which achieve the goal of reducing the dimensionality of the assay readouts and simultaneously identifying individual responses and groups of responses that differ between treatment groups or over time.

9.5 Analyses and data sharing 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. Interim blinded safety and immunogenicity data should not be shared outside of the SMB, HVTN 137 PSRT, the protocol team leadership, the HVTN Executive Management Team, the study product developer, and the study sponsor and/or its designee(s) for their regulatory reporting unless approved by the protocol leadership and the HVTN leadership.

9.5.1 Safety analyses

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 137 PSRT.

9.5.2 Immunogenicity analyses

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 the primary immunogenicity endpoint (neutralizing antibody response) reaches the aforementioned threshold. The Laboratory Center will review the analysis report prior to distribution to the protocol chairs, DAIDS, study product developer, and other key HVTN members and investigators. Reports for distribution or presentation should use PubIDs and not participant ID (PTIDs) for individual responses. 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. SAFETY TABLES AND FIGURES

10.1 List of Tables

SMB reports and Safety FSRs include the following tables.

- Enrollment Report
- Demographics and Study Product Administration Frequencies
- Overall Protocol Status
- Maximum Local Reactogenicity Summary
- Maximum Systemic Reactogenicity Summary
- Adverse Events by Body System and Severity – By Decreasing Frequency
- Adverse Events by Preferred Term and Severity – By Decreasing Frequency – Includes Severe, Life-Threatening or Death Events Only
- Adverse Events by Preferred Term and Severity – By Decreasing Frequency – Includes All Severities
- Adverse Events by Preferred Term and Severity – By Decreasing Frequency – Includes Related Events Only
- Adverse Events by Preferred Term and Relationship to Study Product – By Decreasing Frequency – Includes Events of Any Relationship
- Expedited Adverse Experiences (EAEs) Reported to the Regulatory Support Center (RSC) Listing
- Adverse Events of Special Interest (AESI) Listing
- Medically Attended Adverse Events (MAAE) Listing
- Pregnancy Listing
- HIV Infection Results from Lab and Reported by Site Listing
- Study Product Administration Errors Listing

Safety FSRs include the following additional tables.

- Social Impact Summary
- Local Lab Value Summary Statistics
- Listing of Adverse Events by Treatment Arm, Participant, and Onset Date
- End of Study Diagnostic ELISA Testing Results Listing
- Local Laboratory Values Meeting Grade 1 AE Criteria or Above Listing

10.2 Participant Listings

The following listings of participant-level data are included in SMB reports only.

- Discontinuations
- Severe or Life-Threatening Local Reactogenicities

- Severe or Life-Threatening Systemic Reactogenicities
- Moderate, Severe or Life-Threatening Erythema and Induration
- Severe, Life-Threatening, or Fatal Adverse Events
- Related Adverse Events

10.3 List of Graphs

SMB reports and Safety FSRs include the following graphs.

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

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

11.1 Intracellular Cytokine Staining (ICS)

HIV-1-specific CD4+ and CD8+ T-cell responses will be measured by intracellular cytokine staining as previously described (De Rosa, 2012 & Horton, 2007). CD4+ and CD8+ T cells expressing IFN-γ and/or IL-2 as well as CD4+ T cells expressing IFN-γ and/or IL-2 and/or CD154 will be analyzed. A validated 27-color panel (experiment assay ID 117, analysis plan 051 for Part A and experiment assay ID 123, analysis plan 059 for Part B) will be used to evaluate BG505 gp41 and BG505 gp120 at the following visits:

- Part A: visit 6 (Month 0.5; 2 weeks post 1st vaccination), visit 9 (Month 2.5; 2 weeks post 2nd vaccination), and possibly visit 14 (Month 8).
- Part A with optional second boost: visit 6 (Month 0.5; 2 weeks post 1st vaccination), visit 9 (Month 2.5; 2 weeks post 2nd vaccination), visit 13 (Month 6.5), and possibly visits 15 (Month 12) and 16 (Month 18).
- Parts B and C: visit 9 (Month 2.5; 2 weeks post 2nd vaccination), visit 13 (Month 6.5), and possibly visits 15 (Month 12) and 16 (Month 18).

Several criteria are used to determine if data from an assay are acceptable and can 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 is not, a new specimen for that participant at that time point will be thawed for testing. If the PBMC viability of the second thawed aliquot is below this threshold, the ICS assay will not be performed and no data will be reported to the statistical center for the participant-time point. For the negative control acceptance criteria in Part A only, if the average cytokine response for the negative control wells is above 0.1% for either the CD4+ or CD8+ T cells, the sample will be retested. If the retested results are above 0.1%, the data will be excluded from analysis; otherwise, the retest data will be used.

The total numbers of CD4+ and CD8+ T cells must also have exceeded certain thresholds. If the number of CD4+ T cells is < 10,000 or the number of CD8+ T cells is < 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 will be filtered. If both negative control replicates fail for number of T cells, the sample will be retested. If one negative control replicate fails for number of T cells, the negative control replicate with sufficient cells will be used.

Response magnitudes of CD4+ and CD8+ T cells to Total Env will be calculated as the sum of BG505 gp41 and BG505 gp120. Responder to Any Env will be taken as the maximum response of BG505 gp41 and BG505 gp120. All pools are included regardless of positive or negative responder status, and negative magnitudes are not censored at 0.

To assess positivity for a peptide pool within a T-cell subset, a two-by-two contingency table will be constructed comparing the HIV-1 peptide stimulated and negative control data. The four entries in each table are the number of cells positive for IFN-γ and/or IL-2 and the number of cells negative for IFN-γ and/or IL-2 for both the stimulated and the negative control data. If both negative control replicates are included, then the average number of total cells and the average number of positive cells are used. A one-sided Fisher's exact test is applied to the table, testing whether the number of cytokine-producing cells for the stimulated data is equal to that for the negative control data. Since multiple individual tests (for each peptide pool) are conducted simultaneously, a multiplicity adjustment will be made to the individual peptide pool p-values using the Bonferroni-Holm adjustment method. If the adjusted p-value for a peptide pool is ≤ 0.00001, the response to the peptide pool for the T-cell subset is considered positive. Because the sample sizes (i.e., total cell counts for the T-cell subset) are 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 is chosen stringently (≤ 0.00001). If at least one peptide pool for a specific HIV-1 protein is positive, then the overall response to the protein is considered positive. If any peptide pool is positive for a T-cell subset, then the overall response for that T-cell subset is considered positive. Data from the placebo groups in Part A (P1 and P2) and Part B (P3, P4, P5, and P6) will be pooled as Group C1 and Group C2 respectively.

Naïve (N), Central Memory (CM), Effector Memory (EM), and Terminally Differentiated (TEMRA) memory cell subsets will be analyzed in the final report as a percent of CD4 and CD8 T-cells, as well as a relative percent of cytokine-expressing IFN-γ and/or IL-2 subsets for positive responders only. Relative percentages of each memory cell subset within a participant will sum to 100%.

11.1.1 List of Tables

- Response rate table by T cell subset, antigen, cytokine, 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, antigen, cytokine, 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, antigen, cytokine, visit number (visit month), and treatment group.Table of comparison of response rates between treatment groups by isotype, antigen, and visit among positive responders.
- Table of comparison of response magnitudes between treatment groups by isotype, antigen, and visit among all responders.

11.1.2 List of Graphs

- Boxplots of background-subtracted percent of cells expressing IFN-γ and/or IL-2 by T cell subset, antigen, cytokine, visit number (visit month), and treatment group.
- Spaghetti plots of background-subtracted percent of cells expressing IFN-γ and/or IL-2 over time, by T cell subset, antigen, cytokine, and treatment group.
- Spaghetti plots of background-subtracted percent of T cells expressing naïve and memory cell subsets, by T cell subset, memory cell subset, and treatment group (4 rows of memory cell subsets per page).
- Spaghetti plots of relative percent of naïve and memory cell subsets out of positive T cells expressing IFN-γ and/or IL-2, by T cell subset, memory cell subset, and treatment group (4 rows of memory cell subsets per page).
- Reverse CDFs of background-subtracted percent of cells expressing IFN-γ and/or IL-2 by T cell subset, antigen, cytokine, and treatment group.
- Scatterplot of background-subtracted percent of cells expressing IFN-γ and/or IL-2 at Month 2.5 vs. prime interval, by T cell subset, antigen for Part A
- Scatterplot of background-subtracted percent of cells expressing IFN-γ and/or IL-2 at Month 6.5 vs. prime interval, by T cell subset, antigen for Part A
- Scatterplot of background-subtracted percent of cells expressing IFN-γ and/or IL-2 at Month 6.5 vs. boost interval, by T cell subset, antigen

11.2 B-Cell Phenotyping (BCP)

HIV-1 Env-specific B cells induced by vaccination will be identified and characterized using fluorescently-labelled recombinant Env proteins (BG505-SOSIP.664 and BG505-SOSIP.S241N) in combination with a flow cytometry antibody panel at the following visits:

- Part A: visit 2 (baseline, pre vaccination), visit 6 (Month 0.5; 2 weeks post 1st vaccination), visit 9 (Month 2.5; 2 weeks post 2nd vaccination), and possibly visit 14 (Month 8).
- Part A with optional second boost: visit 6 (Month 0.5; 2 weeks post 1st vaccination), visit 9 (Month 2.5; 2 weeks post 2nd vaccination), visit 13 (Month 6.5), and possibly visits 15 (Month 12) and 16 (Month 18).
- Parts B and C: visit 2 (baseline pre vaccination), visit 9 (Month 2.5; 2 weeks post 2nd vaccination), visit 13 (Month 6.5), and possibly visits 15 (Month 12) and 16 (Month 18).

Live total B cells are identified using doublet exclusion, lymphocyte scatter profile, viability dye, and the following lineage markers: negative for CD3, CD56 and CD14, and positive for CD19 and CD20. IgG+ B cells are further gated on IgD negative and IgG+.

All data reported are the raw frequencies with no background or baseline subtraction or adjustments and are listed below:

- The % BG505-SOSIP.664+ of total B cells equals the frequency of double-positive BG505-SOSIP.664 gp140+ B cells out of total B cells.
- The % BG505-SOSIP.S241N+ of total B cells equals the frequency of double-positive BG505-SOSIP.664 gp140+ B cells that are negative for the BG505-SOSIP.664 mutant (BG505-SOSIP.S241N) out of total B cells. (Part A only)
- The % BG505-SOSIP.664+ of IgG+ B cells equals the frequency of double-positive BG505-SOSIP.664 gp140+ IgG+ B cells out of IgG+ B cells.
- The % BG505-SOSIP.S241N+ of IgG+ B cells equals the frequency of double-positive BG505-SOSIP.664 gp140+ IgG+ B cells that are negative for the BG505-SOSIP.664 mutant (BG505-SOSIP.S241N) out of IgG+ B cells. (Part A only)
- The % BG505-SOSIP.664+ IgG+ of total B cells equals the frequency of double-positive BG505-SOSIP.664 gp140+ IgG+ B cells out of total B cells. (Part B only)
- The % BG505-SOSIP.664+ IgG+ of memory B cells equals the frequency of double-positive BG505-SOSIP.664 gp140+ IgG+ B cells out of memory B cells. (Part B only)
- The % BG505-SOSIP.664+ IgM+ of total B cells equals the frequency of double-positive BG505-SOSIP.664 gp140+ IgM+ B cells out of total B cells. (Part B only)
- The % BG505-SOSIP.664+ IgM+ of memory B cells equals the frequency of double-positive BG505-SOSIP.664 gp140+ IgM+ B cells out of memory B cells. (Part B only)
- The % BG505-SOSIP.664+ IgM+ IgD+ of total B cells equals the frequency of double-positive BG505-SOSIP.664 gp140+ IgM+ IgD+ B cells out of total B cells. (Part B only)
- The % BG505-SOSIP.664+ IgA+ of total B cells equals the frequency of double-positive BG505-SOSIP.664 gp140+ IgA+ B cells out of total B cells. (Part B only)
- The % BG505-SOSIP.664+ IgA+ of memory B cells equals the frequency of double-positive BG505-SOSIP.664 gp140+ IgA+ B cells out of memory B cells. (Part B only)

- The % Plasmablasts of B cells equals the frequency of plasmablast cells out of B cells. (Part B only)
- The % BG505-SOSIP.664+ plasmablast cells of total plasmablasts equals the frequency of double-positive BG505-SOSIP.664 gp140+ plasmablast cells out of total plasmablast cells. (Part B only)

To assess positivity for the detection of SOSIP.664 gp140+ and BG505-SOSIP.S241N gp140+ B cells, Fisher's exact test will be used: A two-by-two contingency table is constructed comparing the post-vaccination and baseline (visit 2) data. The four entries in each table were the numbers of Env-specific B cells and non-Env-specific B cells after vaccination and at baseline. A one-sided Fisher's exact test is applied to the table, testing whether the number of Env-specific B cells for the post-vaccination data is equal to that for the data from baseline, versus an alternative hypothesis that it is greater. Because the sample sizes (i.e., total cell counts for the B-cell subset) are 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 p-value significant threshold is chosen stringently ($\leq$ 0.00001). Data from the placebo groups in Part A (P1 and P2), Part B (P3, P4, P5, and P6), and Part C (P7 and P8) will be pooled as Group C1, Group C2, and Group C3 respectively.

11.2.1 List of Tables

- Response rate table by B cell subset, visit number (visit month), and treatment group.
- Summary statistics (minimum, 25th percentile, median, 75th percentile, maximum, mean, standard deviation) among all participants, by B cell subset, visit number (visit month), and treatment group.
- Summary statistics (minimum, 25th percentile, median, 75th percentile, maximum, mean, standard deviation) among positive responders by B cell subset, visit number (visit month), and treatment group.
- Table of comparison of Env-specific B cells response rates between treatment groups by B cell subset and visit
- Table of comparison of Env-specific B cell response magnitudes between treatment groups by B cell subset and visit among all responders and positive responders only.

11.2.2 List of Graphs

- Boxplots of % HIV-1 Env-specific B cell by protein, visit number (visit month), and treatment group.
- Spaghetti plots of % HIV-1 Env-specific B cell by protein and treatment group.
- Reverse CDFs of % HIV-1 Env-specific B cell by protein, visit number (visit month), and treatment group.
- Scatterplot of % HIV-1 Env-specific B cell at Month 2.5 vs. prime interval, by protein for Part A
- Scatterplot of % HIV-1 Env-specific B cell at Month 6.5 vs. prime interval, by protein for Part A

- Scatterplot of % HIV-1 Env-specific B cell at Month 6.5 vs. boost interval, by protein for Part A

11.3 Neutralizing Antibody (NAb)

Neutralizing antibodies (NAb) against the autologous strain as well as heterologous tier 1A and tier 2 strains of HIV-1 will be measured in TZM-bl cells as a function of reduction in Tat-induced luciferase (Luc) reporter gene expression after a single round of infection with molecularly cloned, Env-pseudotyped viruses. TZM-bl (also called JC57BL-13) is a HeLa cell clone that was engineered to express CD4 and CCR5 and to contain integrated reported genes for firefly luciferase and E. coli β-galactosidase under control of an HIV-1 long terminal repeat (LTR). The cells are highly permissive to infection by most strains of HIV, including primary HIV-1 isolates and molecularly cloned Env-pseudotyped viruses. DEAE-dextran is used in the medium during neutralization assays to enhance infectivity. Expression of the reporter genes is induced in trans by viral Tat protein soon after infection. Luciferase activity is quantified by luminescence and is directly proportional to the number of infectious virus particles present in the initial inoculum. The assay is performed in 96-well culture plates for high throughput capacity. Use of a clonal cell population provides enhanced precision and uniformity. The assay has been formally optimized and validated for single-round infection with either uncloned or molecularly cloned Env-pseudotyped viruses produced by transfection in 293T cells.

Assays will be performed using cryopreserved serum samples starting at a 1:10 dilution. Titers will be defined as the serum dilution that reduces relative luminescence units (RLU) by 50% and 80% relative to the RLU in the virus control wells after subtraction of background RLU in control wells (ID50 and ID80). Neutralization ID50 and ID80 titers will be measured against BG505/T332N (autologous strain) and MW965.26 (heterologous tier 1A strain). Any samples that are positive (ID50>10) against BG505/T332N will be assayed against the 9 virus global panel of heterologous tier 2 viruses: 246F3, Ce1176, CNE55, X1632, Ce0217, BJOX2000, 25710, TRO11, and CH119.

A response is considered positive if the neutralization titer is above 10, the starting dilution, which is the limit of detection of the assay. Magnitude of response will be measured by the natural log of the ID50 and ID80 titer. An aggregate measure of response 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 tier 2. If a titer is left censored, half the left censor limit will be used as the titer value.

For Part A of the study, neutralization titers will be measured on samples from the primary immunogenicity time point, week 10 (visit 9 – 2 weeks post 2nd vaccination) and as applicable, week 26 (visit 13 – 2 weeks post optional 2nd boost). Specimens from other time points may also be analyzed at the discretion of the PI, which may be contingent on the results of the primary immunogenicity time point. Those time points may include week 0 (visit 2 – baseline), week 2 (visit 6 – 2 weeks post 1st vaccination), and/or week 32 (visit 14 – 6 months post 2nd vaccination). For those who consented to the optional second boost, additional time points may include week 0 (visit 2 – baseline), week 2 (visit 6 – 2 weeks post 1st vaccination, week 24 (visit 11 – optional 2nd boost visit), week 52 (visit 15 – 6 months post optional 2nd vaccination), and/or week 78 (visit 16, 12 months post optional 2nd vaccination). Data from the two placebo groups in Part A (P1 and P2) will be pooled as Group C1.

For Part B of the study, neutralization titers will be measured on samples from the primary immunogenicity time point, week 10 (visit 9 – 2 weeks post 2nd vaccination) and week 26 (visit 13 – 2 weeks post 3rd vaccination). Specimens from other time points may also be analyzed at the discretion of the PI, which may be contingent on the results of the primary immunogenicity time point. Those time points may include week 0 (visit 2 – baseline), week 52 (visit 15 – 6 months post 3rd vaccination), and/or week 78 (visit 16, 12 months post optional 2nd vaccination). Data from the four placebo groups in Part B (P3, P4, P5, and P6) will be pooled as Group C2.

11.3.1 List of Tables

- Response rate table by cell type, isolate, visit number (visit month), and treatment group.
- Summary statistics for NAb titers (minimum, 25th percentile, median, 75th percentile, maximum, mean, standard deviation) among all participants, by cell type, isolate, visit number (visit month), and treatment group.
- Summary statistics for NAb titers (minimum, 25th percentile, median, 75th percentile, maximum, mean, standard deviation) among positive responders by cell type, isolate, visit number (visit month), and treatment group.
- Listing of neutralizing antibody titers among positive responders.

11.3.2 List of Graphs

- Boxplots of neutralizing antibody titers by cell type, isolate, visit, and treatment group.
- Spaghetti plots of neutralizing antibody titers over time by cell type, isolate, and treatment group.
- Reverse CDFs of neutralizing antibody titers by cell type, isolate, visit and treatment group.
- Scatterplot of neutralizing antibody titers at Month 2.5 vs. prime interval, by cell type, isolate for Part A
- Scatterplot of neutralizing antibody titers at Month 6.5 vs. prime interval, by cell type, isolate for Part A
- Scatterplot of neutralizing antibody titers at Month 6.5 vs. boost interval, by cell type, isolate for Part A

11.4 Binding Antibody Multiplex Assay (BAMA)

Serum HIV-1-specific total binding IgG antibody responses will be measured at the 1:50 dilution against the gp41 and BG505 SOSIP.664 AviB antigens on samples collected from Part A and Part B's participants at week 0 (visit 2 – baseline), week 10 (visit 9 – 2 weeks post 2nd vaccination), and if applicable, week 26 (visit 13 – 2 weeks post optional 2nd boost or 3rd vaccination), week 32 (visit 14), week 52 (visit 15), and week 78 (visit 16). Data from the placebo groups in Part A (P1 and P2) and Part B (P3, P4, P5, P6) will be pooled as Group C1 and Group C2 respectively. BAMA testing for Part C will include the Trimer 4571 antigen and will be performed on all available participant serum samples collected from Part C study arm groups 7 and 8 at visits 2 (month 0, baseline), 9 (month 2.5, 2 weeks post second vaccination), 13 (month 6.5, 2 weeks post third vaccination), 15 (month 12, 6 months post third vaccination), and 16 (month 18, 12 months post third vaccination).

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 or equal to 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 (AUC), in which case geometric mean titers between treatment groups at each time point will be compared using two-sample T-test comparing the means of log AUC.

In addition, HIV-1-specific total binding IgA antibodies, binding to IgG subclasses (IgG1, IgG2, IgG3, and IgG4), and Fc-mediated antibody functions (e.g. fragment crystallizable receptor (FcR) array) may also be assessed. Specimens from other timepoints may also be assayed based on results of the initial assay. Other exploratory analyses will examine HIV-1 Env IgG and IgA antibody titers in mucosal (rectal and vaginal) tissues and secretions. Mucosal binding antibody magnitude is quantified in terms of specific activity (SA). SA is defined as MAX(0.0002,MFI-Blank*dilution/total antibody concentration).

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.

11.4.1 List of Tables

- Response rate table by isotype, antigen, visit number (visit month), and treatment group.
- Summary statistics of MFI* values (minimum, 25th percentile, median, 75th percentile, maximum, mean, standard deviation) among all participants, by isotype, antigen, visit number (visit month), and treatment group.
- Summary statistics of MFI* values (minimum, 25th percentile, median, 75th percentile, maximum, mean, standard deviation) among positive responders by isotype, antigen, visit number (visit month), and treatment group.
- Summary statistics of specific activity (minimum, 25th percentile, median, 75th percentile, maximum, mean, standard deviation) among all participants, by isotype, antigen, visit number (visit month), and treatment group. (Mucosal specimens only)
- Summary statistics of specific activity (minimum, 25th percentile, median, 75th percentile, maximum, mean, standard deviation) among positive responders by isotype, antigen, visit number (visit month), and treatment group. (Mucosal specimens only)
- Summary statistics of binding antibody titers (AUC) (minimum, 25th percentile, median, 75th percentile, maximum, mean, standard deviation) among all participants, by isotype, antigen, visit number (visit month), and treatment group. The 95% CI will be calculated based on a normal approximation for log-transformed AUC.
- Summary statistics of binding antibody titers (AUC) (25th percentile, median, 75th percentile, maximum, mean, standard deviation) among positive responders, by isotype, antigen, visit number (visit month), and treatment group. The 95% CI will be calculated based on a normal approximation for log-transformed AUC.
- Tables of comparison of geometric means of binding antibody titers (AUC) between treatment groups by isotype, antigen, and visit among all participants and among positive responders.

11.4.2 List of Graphs

- Boxplots of MFI* values by isotype, antigen, visit, and treatment group.
- Boxplots of binding antibody titers (AUC) by isotype, antigen, visit, and treatment group.
- Boxplots of specific activity by isotype, antigen, visit, and treatment group. (Mucosal specimens only)
- Magnitude-Breadth curves grouped by individual treatment arms and pooled treatment arms
- Scatterplot of binding antibody titers at Month 2.5 vs. prime interval, by isotype, antigen for Part A
- Scatterplot of binding antibody titers at Month 6.5 vs. prime interval, by isotype, antigen for Part A
- Scatterplot of binding antibody titers at Month 6.5 vs. boost interval, by isotype, antigen for Part A

11.5 Trucount Assay

The Trucount® pan-lineage panel is designed to provide counts for the major lineages found among leukocytes in blood. It is used on whole blood with Trucount® beads to allow for determination of absolute cell concentrations. There is a fixed number of beads within each sample at the time of staining and the number is determined by the Trucount® lot number. Whole blood stained for Trucount® is analyzed on an LSR II cytometer. Unlike a functional assay, there are no functional negative or positive controls.

The primary Trucount readout endpoints are cell subset concentrations (cells/µl blood) and cell subset percentages. Subset concentrations and cell percentages of interest include total PBMCs, neutrophils, CD11b+ neutrophils, CD16+ neutrophils, total monocytes, CD86+ monocytes, classical monocytes, CD86+ classical monocytes, non-classical monocytes, CD86+ non-classical monocytes, intermediate monocytes, CD86+ intermediate monocytes, lymphocytes, CD3+ T cells, CD4+ T cells and their activated (CD38+, CD38+DR+ double positive, CD69+, HLA-DR+) subsets, CD4- CD8- T cells and their activated (CD38+, CD38+DR+ double positive, CD69+, HLA-DR+) subsets, CD8+ T cells and their activated (CD38+, CD38+DR+ double positive, CD69+, HLA-DR+) subsets, total NK cells, 16+56+ NK cells, 16+56- NK cells, 16-56- NK cells, dendritic cells and their CD86+ subset, myeloid dendritic cells (mDC), CD86+ mDCs, plasmacytoid DC (pDC), CD86+ pDC, CD19+ B cells, and CD20loCD38hi plasmablasts CD19+ B cells and their CD86+ subset. Cell concentrations are calculated based on the relative proportion of Trucount® beads in each tube, and percentages are calculated based on parent populations. Concentrations are calculated taking into account the dilution factor introduced by the ACD liquid anticoagulant.

Trucount® assays are planned for up to 88 participants from Part B and 22 participants from Part C. Assay data over time at innate immune time points, Day 0, 1, 3, 7, 56 and 168 are available after the first vaccination. The blood samples at innate immune time point, Day 0, were collected prior to the vaccination. The cell concentrations at Day 0, therefore, will be treated as baseline measures for the cell concentrations at the later innate immune time points (Days 1, 3, 7, 56, 168) after the same vaccination.

The cell concentrations will be presented on the log-scale graphically and be log-transformed prior to statistical analysis. Cell percentages will be presented and analyzed on the linear scale.

The following comparisons of innate immune responses over time within each Group or between Groups within a time point will be done:

- The innate immune responses at Days 1, 3, 7, 56, and 168 will be compared to the innate immune responses at Day 0 within each group. The comparisons will be done using a linear mixed effect (LME) model with random effect for subject and fixed effects for Group and Day (modeled as a factor variable), with an interaction between Day and Group. Both log-cell concentrations and cell percentages will be evaluated—in separate mixed effects models. The assumptions of the LME will be evaluated using standard diagnostics, e.g. by plotting residuals.

11.5.1 List of Tables

- Summary statistics (i.e., min, mean, median, max) of the cell concentrations by innate immune time points (Days 0, 1, 3, 7, 56, 168) and treatment group
- Summary statistics (i.e., min, mean, median, max) of the cell percentages by innate immune time points (Days 0, 1, 3, 7, 56, 168) and treatment group

11.5.2 List of Graphs

- Plots of individuals' cell concentrations over time by treatment with the median for each treatment group superimposed on the individuals' profiles. The cell concentrations will be plotted on the log scale.
- Plots of individuals' cell percentages over time by treatment group with the median for each treatment group superimposed on the individuals' profiles.

11.6 Electron Microscopy Polyclonal Epitope Mapping (EMPEM)

Polyclonal antibodies (IgG) are isolated from individual (participant) blood samples using commercial Fc-affinity purification resins. The antibodies are enzymatically digested into the fragment antigen binding (Fab) components with papain, and incubated with soluble, HIV Env trimer proteins, matched to the immunogen used in the study (BG505 SOSIP.664 or Trimer 4571). The complex is purified by size-exclusion chromatography, adsorbed onto electron microscopy (EM) grids, stained, and imaged. Individual protein complex particles are extracted from the images and subjected to averaging and classification in 2D and 3D space. 3D focused classification methods are used to evaluate each of the defined epitopes on the surface of Env, both in qualitative and quantitative (the magnitude of the given epitope response) terms. The end results are 3D EM maps that are matched to known structures of Env in complex with antibodies and each polyclonal Fab specificity is assigned a final epitope label based on overlap with known structure(s).

The non-quantitative results are graphical and illustrative representations of all unique Env epitopes detected, as a function of individual and timepoint. The quantitative results are per-epitope EMPEM magnitudes, calculated as a function of the number of particles containing a given polyclonal Fab relative to the total number of particles evaluated in 3D space. The per-epitope magnitude value is on a scale of 0-3 to account for trimeric Env having up to 3 copies of a given epitope. It is analogous to stoichiometry (i.e. a value of 3 suggests that all particles had all three copies of the same epitope fully occupied by polyclonal Fab). Note that protomers are an observed phenotype (trimer dissociation/disassembly) and not a unique antibody. These epitopes

cannot be accurately mapped and as such have only non-quantitative results. Summaries of response are number of epitope-specific antibodies detected, the relative abundance of each epitope-specific antibody in a given sample, and number of epitope-specific antibodies excluding responses detected to the gp41-base. Serum specimens are collected from participants enrolled in Parts A, B, and C at visit 9 (week 10) and visit 13 (week 26). The assay is further described in Bianchi et al. (2018), Turner et al. (2023), and Hahn et al. (2024).

The epitopes assignments are: gp41-base, V1V2V3, gp41-GH, C3V5, CD4bs, gp120-GH, gp120-gp120, and gp41-FP (GH: glycan hole, FP: fusion peptide, CD4bs: CD4 binding site). If antibody-induced trimer disassembly is detected during 2D classification, the sample is also assigned a label: protomers.

Examples of EMPEM-derived 3D maps and the location of each epitope with respect to HIV Env ectodomain are shown in Figure 1.

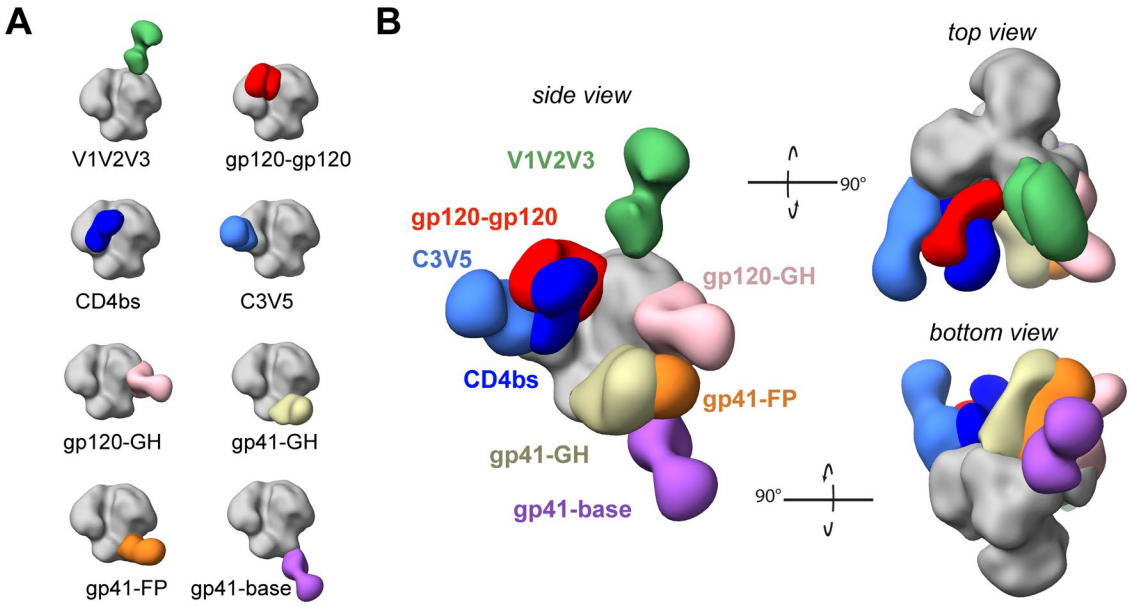

Figure 1 Visualization of EMPEM-derived 3D maps. A) Representative maps of antibody-trimer complexes for all defined EMPEM epitopes. B) Overlay of antibodies to the 8 defined EMPEM epitopes.

11.6.1 List of Tables

- Response rate table by epitope, visit number (visit month), and treatment group.
- Response rate table of antibody-induced trimer disassembly (detected protomers) by visit number (visit month) and treatment group.
- Summary statistics (mean, median) of epitope-specific antibodies (with and without responses directed to the base, excluding protomers) by treatment group and visit number (visit month)

11.6.2 List of Graphs

- Box and bar plots show both response rates and the distribution of magnitude by epitope where the boxplot is drawn around only positive responders

- Box and bar plots show the rate of any positive response across all epitopes, and the distribution of the number of epitope-specific responses, where the boxplot is drawn around those positive for at least one epitope. This is shown with and without counting the gp41 base epitope
- Box and bar plots show the rate of any positive response across all epitopes and the distribution of total magnitude, with or without including the gp41-base, where the boxplot is drawn around only positive responders
- The above box and barplots are repeated, where boxplots show the distribution of response among all participants, not just among positive responders
- Bar plots illustrate the percentage of participants with protomers detected
- A participant-level plot shows which epitopes were positive for each participant and treatment group

11.7 Infected cells antibody binding assay (ICABA)

The ICABA assay is conducted to address the exploratory objective to evaluate immunogenicity of the vaccine regimens for Parts B and C. ICABA will detect the presence of antibodies binding to the surface of HIV-1 infected cells in the serum of participants using Infectious Molecular Clone (IMC)-infected target cells.

For the analysis, ICABA assays will be performed on participants' sera from Parts B and C for v2 (M0) and v13 (M6.5). In the ICABA, eight-six participant samples can be tested per plate. All participant samples from this study can be run in two assays of two plates total (one plate of mock and one plate of infected cells).

CEM.CCR5.NKR cells infected with HIV-1 IMC BG505 and mock-infected cells will be tested in a 96-well plate for the ICABA. Participant sera in addition to controls will be incubated with the IMC-infected cells and the presence of bound antibodies will be detected on the surface of the cells using flow cytometry. If the infected cell mock- and baseline-subtracted %IgG+ at 1:100 dilution is >5%, it is considered positive.

11.7.1 List of Tables

- Response rate table by visit number (visit month) and treatment arm
- Summary statistics (i.e., min, mean, median, max) among all participants by visit and treatment arm
- Summary statistics (i.e., min, mean, median, max) among positive responders by visit and treatment arm

11.7.2 List of Graphs

- Boxplot of mock- and baseline-subtracted %IgG+ by visit and treatment arm.
- Spaghetti plot of mock- and baseline-subtracted %IgG+ over time by treatment arm.

11.8 Antibody-Dependent Cell-mediated Cytotoxicity (ADCC)

The established GranToxiLux (GTL) and Luciferase assays will be used to assess response rates and magnitude of serum ADCC. ADCC assays will be performed on participants' sera from Parts B and C for v2 (M0), v9 (M2.5), and v13 (M6.5).

CEM.CCR5.NKR cells infected with HIV-1 IMC BG505 will be tested in a 96-well plate for the Luciferase ADCC assay. Participant sera in addition to controls will be incubated with the IMC-infected cells, and ADCC will be detected through the use of Viviren luminescence. To assess breadth, CEM.CCR5.NKR cells infected with HIV-1 CAP8, CH058, SUMA, and WITO, and three additional IMCs (to be determined), individually, will be tested in a 96-well plate for the Luciferase assay.

CEM.CCR5.NKR cells coated with Clade A BG505 gp120 will be tested in a 96-well plate for the GranToxilux (GTL) assay. Participant sera in addition to controls will be incubated with the HIV protein, and ADCC will be detected through the use of granzyme B substrate (GTL).

The following evaluation criteria will be employed for these assays:

i) Luciferase: The data from the luciferase assay is semiquantitative. The peak of specific killing and area under the curve (AUC) will be compared between the placebo and vaccine groups.

ii) GTL: The data from the assay is quantitative. The frequency of positive responses (% responders) peak ADCC, and area under the curve (AUC) will be tabulated to assess differences in ADCC response between the treatment groups.

Positivity criteria are as follows:

i) Luciferase: peak % loss Luciferase activity ≥10%, after background subtraction of baseline response and positivity at one of the first two dilutions.

ii) GTL: peak activity ≥ 8%

11.8.1 List of Tables

Luciferase:

- Response rate of peak background subtracted percent killing or loss luciferase activity by antigen, visit number (visit month), and treatment arm
- Summary statistics (i.e., min, mean, median, max) of peak background subtracted percent killing or loss luciferase activity among all participants by antigen, visit, and treatment arm
- Summary statistics (i.e., min, mean, median, max) of peak background subtracted percent killing or loss luciferase activity among positive responders by antigen, visit, and treatment arm
- Summary statistics (i.e., min, mean, median, max) of pAUC of background subtracted percent killing among all participants by treatment group, protein, and visit
- Summary statistics (i.e., min, mean, median, max) of pAUC of background subtracted percent killing among positive responders by treatment group, protein, and visit
- Response rate comparison using Barnard's exact test

GTL:

- Response rate of peak activity by antigen, visit number (visit month), and treatment arm
- Summary statistics (i.e., min, mean, median, max) of peak activity among all participants by antigen, visit, and treatment arm

- Summary statistics (i.e., min, mean, median, max) of peak activity among positive responders by antigen, visit, and treatment arm
- Summary statistics (i.e., min, mean, median, max) of AUC by treatment group, protein, and visit
- Summary statistics (i.e., min, mean, median, max) of AUC among positive responders by treatment group, protein, and visit

11.8.2 List of Graphs

- Barcharts of GTL peak activity response rates and boxplots of peak activity response magnitudes study visit, treatment group and antigen
- Barcharts of ADCC Luciferase response rates and boxplots of ADCC response magnitudes by study visit, treatment group and antigen
- Barcharts of GTL peak activity response rates and boxplots of AUC by study visit, treatment group and antigen
- Barcharts of ADCC Luciferase response rates and boxplots of pAUC of background subtracted percent killing by study visit, treatment group and antigen

11.9 Antibody-Dependent Cellular Phagocytosis (ADCP)

The Antibody Dependent Cellulary Phagocytosis (ADCP) is a qualified assay employing flow cytometric-based technology that measures the ability of antibodies to mediate phagocytosis. Biotinylated antigen conjugated to neutravidin beads are incubated with patient serum or plasma, purified IgG, or monoclonal antibodies. A monocyte cell line (THP-1) is added to the immune complexes and spinoculated at 4ºC followed by incubation at 37ºC. Cells are then analyzed for bead internalization by flow cytometry (bead positive versus bead negative detection).

A phagocytic score is determined based on the ratio of experimental sample to PBS control. Mean phagocytosis score is defined as: (% bead positive for participant x MFI bead positive for participant) / (% bead positive for PBS only control x MFI bead positive for PBS only control).

Positivity calls will be based on two criteria:

i. ADCP score greater than the 95th percentile average of the visit 2 samples.

ii. ADCP score greater than three-fold over the ADCP score of the sample-specific baseline.

Participants' serum samples at timepoints visit 2 (week 0), visit 6 (week 2), visit 9 (week 10) and visit 13 (week 26) will be analyzed by ADCP for Part B. Assays will be performed at visit 2 (week 0) and visit 13 (week 26) for Part C.

11.9.1 List of Tables

- ADCP response rates by treatment group, study visit, and antigen
- Distribution of phagocytosis score among all participants by treatment group, study visit, and antigen
- Distribution of phagocytosis score among positive responders only by treatment group, study visit, and antigen

11.9.2 List of Graphs

- Barcharts of ADCP response rates and boxplots of phagocytosis score by treatment group, study visit, and antigen

12. REFERENCES

1. Agresti A, Coull BA. Approximate is better than "exact" for interval estimation of binomial proportions. Am Stat. 1998;52(2):119-26.
2. Dunnett C. A multiple comparisons procedure for comparing several treatments with a control. Journal of the American Statistical Association. 1995;50:1096-121.
3. Brookes ST, Whitely E, Egger M, Smith GD, Mulheran PA, Peters TJ. Subgroup analyses in randomized trials: risks of subgroup-specific analyses; power and sample size for the interaction test. J Clin Epidemiol. 2004;57(3):229-36.
4. Lachenbruch PA. Comparisons of two-part models with competitors. Stat Med. 2001;20(8):1215-34.
5. Lynch RM, Boritz E, Coates EE, DeZure A, Madden P, Costner P, et al. Virologic effects of broadly neutralizing antibody VRC01 administration during chronic HIV-1 infection. Sci Transl Med. 2015;7(319):319ra206.
6. Hughes JP. Mixed effects models with censored data with application to HIV RNA levels. Biometrics. 1999;55(2):625-9.
7. Rotnitzky A, Robins J. Analysis of semi-parametric regression models with non-ignorable non-response. Stat Med. 1997;16(1-3):81-102.
8. Rotnitzky A, Robins JM. Semiparametric regression estimation in the presence of dependent censoring. Biometrika. 1995;82(4):805-20.
9. Gilbert PB, Sato A, Sun X, Mehrotra DV. Efficient and robust method for comparing the immunogenicity of candidate vaccines in randomized clinical trials. Vaccine. 2009;27(3):396-401.
10. James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning with applications in R. New York: R. Springer; 2013.
11. DeCamp A, Hraber P, Bailer RT, Seaman MS, Ochsenbauer C, Kappes J, et al. Global panel of HIV-1 Env reference strains for standardized assessments of vaccine-elicited neutralizing antibodies. J Virol. 2014;88(5):2489-507.
12. Huang Y, Gilbert P, Montefiori D, Self S. Simultaneous evaluation of the magnitude and breadth of a left- and right-censored multivariate response, with application to HIV vaccine development. Statistics in Biopharmaceutical Research. 2009;1:81-91.
13. Finak G, McDavid A, Chattopadhyay P, Dominguez M, De RS, Roederer M, et al. Mixture models for single-cell assays with applications to vaccine studies. Biostatistics. 2014;15(1):87-101.
14. 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.

15. van der Laan MJ, Polley EC, Hubbard AE. Super learner. Stat Appl Genet Mol Biol. 2007;6:Article25.

16. De Rosa, Stephen C, Donald K. Carter, and M. Juliana McElrath. 2012. "OMIP014: Validated Multifunctional Characterization of Antigen Specific Human T Cells by Intracellular Cytokine Staining." Cytometer 81A (12): 1019–21.

17. Horton, Helen, Evan P. Thomas, Jason A. Stucky, Ian Frank, Zoe Moodie, Yunda Huang, Ya-Lin Chiu, M. Juliana McElrath, and Stephen C De Rosa. 2007. "Optimization and Validation of an 8-Color Intracellular Cytokine Staining (Ics) Assay to Quantify Antigen-Specific T Cells Induced by Vaccination." Journal of Immunology Methods 323 (1): 39–54.

18. Bianchi, Matteo, Hannah L. Turner, Bartek Nogal, Christopher A. Cottrell, David Oyen, Matthias Pauthner, Raiza Bastidas, et al. 2018. "Electron-Microscopy-Based Epitope Mapping Defines Specificities of Polyclonal Antibodies Elicited During HIV-1 BG505 Envelope Trimer Immunization." Immunity 49 (2): 288–300.e8. https://doi.org/https://doi.org/10.1016/j.immuni.2018.07.009.

19. Hahn, William O, K. Rachael Parks, Mingchao Shen, Gabriel Ozorowski, Holly Janes, Lamar Ballweber-Fleming, Amanda S Woodward Davis, et al. 2024. "HIV BG505 SOSIP.664 Trimer with 3M-052-AF/Alum Induces Human Autologous Tier-2 Neutralizing Antibodies." medRxiv. Cold Spring Harbor Laboratory Press. https://doi.org/10.1101/2024.05.08.24306957.

20. Turner, Hannah L., Abigail M. Jackson, Sara T. Richey, Leigh M. Sewall, Aleksandar Antanasijevic, Lars Hangartner, and Andrew B. Ward. 2023. "Protocol for Analyzing Antibody Responses to Glycoprotein Antigens Using Electron-Microscopy-Based Polyclonal Epitope Mapping." STAR Protocols 4 (3): 102476. https://doi.org/https://doi.org/10.1016/j.xpro.2023.102476.

13. CHANGE HISTORY

SAP Version	Date	Modification
1.0	02 March 2020	Initial
2.0	23 November 2020	Added ICS, BCP, and NAb immunogenicity analyses for Part A of the study to Sections 11.1, 11.2, and 11.3.
2.1	23 December 2020	Added BAMA immunogenicity analyses for Part A of the study to Section 11.4.
2.2	07 March 2022	Added Part A optional second boost to Sections 3.1.2, 4, 11.2, 11.2, 11.3, and 11.4.
2.3	25 March 2022	Removed reverse CDFs plot in Section 11.4.2 Added Trucount analysis to Section 11.5
2.4	22 April 2022	Updated contingency table description and removed G505-SOSIP.S241N from Part B BCP analysis in Section 11.2
2.5	13 May 2022	Updated ICS assay color panel for Part B and high background filtering for Part A only in Section 11.1
3.0	25 May 2022	Added exploratory analysis (scatterplot) to inform time between vaccination and visit for Part A to Sections 9.1, 9.4.1, 11.1.2, 11.2.2, 11.3.2, and 11.4.2
4.0	18 January 2024	Added EMPEM analysis to Section 11.6

		Updated to include Part C throughout Added ICABA analysis to Section 11.7 Added ADCC analysis to Section 11.8 Added ADCP analysist to Section 11.9
5.0	03 April 2024	Updated ADCC plots for pAUC and AUC in Section 11.8.2
6.0	29 May 2024	Added ICS memory marker analyses to Section 11.1.2. Added ADCC Luciferase breadth to Section 11.8. Added Part C to BCP (Section 11.2) and ICS (Section 11.1).
7.0	30 September 2024	Updated language in EMPEM (Section 11.6)
8.0	07 January 2025	Corrected BAMA response definition to “greater than or equal to” threshold in Section 11.4