Statistical Analysis Plan

Date: 07 June 2017

Author(s): Chris Gast, PhD

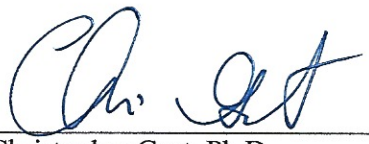
Protocol Number: UAM4a

Protocol Title: A Phase 1, blinded, single-center study to evaluate the safety and immunogenicity of two novel live attenuated serotype 2 oral poliovirus vaccines, derived from a modified Sabin 2 infectious cDNA clone, in healthy adults previously vaccinated with inactivated polio vaccine (IPV)

APPROVAL SIGNATURES

Protocol Number:	UAM4a
SAP Version:	1.1
Version Date:	07 June 2017

By signing below, I confirm that I have reviewed and approve the above version of the Statistical Analysis Plan.



2017/06/07

Christopher Gast, Ph.D.
Biostatistician

Date

Prof. Dr. Pierre Van Damme, Ph.D.
Principal Investigator

Date

Change / Release Log

Date	Version	Changes
2017-05-06	0.1	Initial draft
2017-05-11	0.2	Second draft
2017-05-20	0.3	Third draft
2017-05-25	1.0	Final
2017-06-07	1.1	Addition of mouse gender effect to model-based neurovirulence analyses

Table of Contents

1 List of Abbreviations and Definitions of Terms

1.1 Abbreviations

AE	Adverse event
bpm	Beats per minute
$CCID_{50}$	50% cell culture infective dose
CI	Confidence interval
CRF	Case Report Form
CTCAE	Common Terminology Criteria for Adverse Events
cVDPVs	Circulating vaccine-derived polioviruses
cVDVP2	Circulating vaccine-derived poliovirus type 2
DBP	Diastolic blood pressure
DNA	Deoxyribonucleic acid
eCRF	Electronic Case Report Form
EDC	Electronic Data Capture
GCP	Good Clinical Practice
GMT	Geometric mean titer
GPEI	Global Polio Eradication Initiative
HBsAg	Hepatitis B surface antigen
HCV	Hepatitis C virus
HIV	Human immunodeficiency virus
HR	Heart rate
ICF	Informed Consent Form
ICH	International Conference on Harmonization
IEC	Independent Ethics Committee
IME	Important medical event
IMP	Investigational medicinal product
IPV	Inactivated poliovirus vaccine
IRB	Institutional Review Board
LSLV	Last Subject Last Visit
MedDRA	Medical Dictionary for Regulatory Activities
mOPV2	Monovalent oral poliovirus vaccine type 2
nOPV2	Novel oral poliovirus vaccine type 2
OPV	Oral poliovirus vaccine
PD_{50}	50% paralytic dose
PP	Per-protocol
RNA	Ribonucleic acid
SAE	Serious adverse event
SAGE	Strategic Advisory Group of Experts on immunization
SAP	Statistical Analysis Plan
SBP	Systolic blood pressure
SD	Standard deviation
SmPC	Summary of Product Characteristics
SUSAR	Suspected Unexpected Serious Adverse Reaction

tOPV	Trivalent oral polio vaccine
bOPV	Bivalent oral poliovirus vaccine
TgPVR	Transgenic mice expressing the cell receptor for poliovirus
TMF	Trial Master File
VAPP	Vaccine-associated paralytic poliomyelitis
WHO	World Health Organization
WPV	Wild poliovirus

1.2 Definitions

Seroprotection is defined as type 2-specific antibody titers ≥1:8 and seroprotection rate as the percentage of seroprotected subjects per group.

Seroconversion is defined as a change from seronegative to seropositive and antibody titers of ≥1:8, and in seropositive subjects, as an antibody titer increase of ≥ 4 fold over baseline titers.

See the study protocol for definitions regarding serious adverse events, important medical events, and adverse event severity.

2 Introduction

This document outlines the statistical methods to be implemented for the analysis of the data resulting from Protocol UAM4a, *A Phase 1, blinded, single-center study to evaluate the safety and immunogenicity of two novel live attenuated serotype 2 oral poliovirus vaccines, derived from a modified Sabin 2 infectious cDNA clone, in healthy adults previously vaccinated with inactivated polio vaccine (IPV).* Results of the proposed analyses will be used in the clinical study report for this protocol.

The purpose of this statistical analysis plan (SAP) is to provide specific guidelines for all statistical analyses. All analyses specified in this document will be performed. Any changes will either be reflected in amendments to this plan before the database lock, or documented in the final statistical and clinical study reports. Other analyses that are not included in the SAP may be specified subsequent to its finalization. Such analyses will be described in an addendum as post-hoc and exploratory to the finalized SAP and will be performed if agreed among the participating institutions.

3 Study Overview

3.1 Background and rationale

This is a Phase I contained-use study of the safety and immunogenicity of two novel attenuated serotype 2 oral poliovirus vaccines given to adults who were previously vaccinated with inactivated polio vaccine (IPV). Its purpose is to provide safety and shedding data that will support continued clinical development involving deliberate release studies leading to the licensure of a new type 2 oral poliovirus vaccine, as the serotype 2 Sabin strain oral poliovirus vaccine has been removed from routine immunization strategies worldwide.

3.2 Study objectives

3.2.1 Primary objective

Primary objectives of the study are to assess:

- safety (serious adverse events [SAEs] and severe[1] adverse events [AEs]) of nOPV2 candidate 1 and nOPV2 candidate 2.
- viral shedding following administration of nOPV2 candidate 1 and nOPV2 candidate 2 in all stool samples;

3.2.2 Secondary objectives

Secondary objectives are to assess:

- safety (any solicited and unsolicited AEs, laboratory assessments) of nOPV2 candidate 1 and nOPV2 candidate 2;
- immunogenicity (seroprotection rate, seroconversion rate, median antibody titer (post-vaccination)) of nOPV2 candidate 1 and nOPV2 candidate 2.
- neurovirulence of shed virus (as measured in animal model(s)) in a subset of stool samples of all subjects.

3.2.3 Exploratory objectives

Exploratory objectives are to assess

- immunogenicity (geometric mean titer [GMT]) of nOPV2 candidate 1 and nOPV2 candidate 2;
- genetic stability, including but not limited to the modified regions of shed virus in a subset of stool samples of all subjects.
- viral shedding following nOPV2 candidate 1 or nOPV2 candidate 2 administration in nasopharyngeal swabs of all subjects.

3.3 Study endpoints

3.3.1 Primary endpoints

The following endpoints will be evaluated by group and overall:

Safety:

- incidence, type and causality of SAEs and severe[1] AEs throughout the study period.

[1] List of sev ere AEs as mentioned in the diary cards: fever > 39°C, headache, fatigue, myalgia, arthralgia , paresthesia, anesthesia, paralysis, or gastrointestinal symptoms (nausea, vomiting, diarrhea and/or abdominal pain) that prevent normal activity or any other severe AE that prevents normal activity.

Viral shedding:

- Viral shedding positivity rate (as determined using quantitative PCR) will be assessed at each stool sample collection time point
- Median 50% cell culture infective dose ($CCID_{50}$; titer) of shed virus after viral extraction from stool samples will be assessed at each stool sample collection time point that is positive for type-2 poliovirus via quantitative PCR
- Time-to-cessation of type-2 viral shedding will be assessed
- A combined index of the prevalence, quantity and duration of shedding will be assessed using fixed stool sample collection time points from each subject.

3.3.2 Secondary endpoints

The following safety and immunogenicity endpoints will be evaluated by group and overall.

Safety:

- Incidence, type, causality and severity of solicited AEs for Days 0-7 in Groups 1 and 2;
- Incidence, type, causality and severity of unsolicited AEs throughout the study period in both groups;
- Incidence, causality and description of deviations from normal safety labs at Day 0, Day 7, and Day 28 for Group 1;
- Incidence, causality and description of deviations from normal safety labs at Day 0, Day 7, and Day 28 for Group 2;

Immunogenicity:

- Median titers of type 2 polio antibodies at Days 0 and 28 in Group 1;
- Median titers of type 2 polio antibodies at Days 0 and 28 in Group 2;

- Seroprotection rate of type 2 polio antibodies at Days 0 and 28 in Group 1;
- Seroprotection rate of type 2 polio antibodies at Days 0 and 28 in Group 2;

Seroprotection is defined as type 2-specific antibody titers ≥1:8.

- Seroconversion rate of type 2 polio antibodies at Day 28 for Group 1;
- Seroconversion rate of type 2 polio antibodies at Day 28 for Group 2;

Seroconversion is defined as a change from seronegative to seropositive and antibody titers of ≥1:8, and in seropositive subjects, as an antibody titer increase of ≥ 4 fold over baseline titers.

Viral shedding:

- Neurovirulence of shed virus (as measured in animal model(s)) in a subset of stool samples of all subjects.

3.3.3 Exploratory endpoints

- GMT of type 2 polio antibodies at Days 0 and 28 in Group 1;
- GMT of type 2 polio antibodies at Days 0 and 28 in Group 2;
- Assessment of the genetic stability, including but not limited to the modified regions of shed virus in a subset of stool samples of all volunteers.
- Assessment of nasopharyngeal viral shedding in swabs of all subjects.

3.4 Study design

This will be a single center, blinded study in 30 healthy IPV-only vaccinated adults (age range 18 to 50 years), as follows:

- 15 subjects to receive 1 dose of nOPV2 candidate 1 (Group 1);
- 15 subjects to receive 1 dose of nOPV2 candidate 2 (Group 2).

Subjects who will pass screening but for any reason will drop out before vaccination will be defined as screen-failures and will be replaced.

15 subjects will be evaluated for the 1-dose regimen nOPV2 candidate 1 (Group 1).

15 subjects will be evaluated for the 1-dose regimen nOPV2 candidate 2 (Group 2).

3.5 Schema

The following table provides a summary of the timing of vaccine administration and samples taken from subjects, which will be available for analysis.

Study Group (Country: Belgium)	**N**	**Vaccination**	**Serum Samples**	**Stool Samples (days)****
Group 1: nOPV2, first candidate	15	0	0, 28	Screening (> Day -14), then continuously following vaccination, until no more shedders
Group 2: nOPV2, second candidate	15	0*	0, 28	Screening (> Day -14), then continuously following vaccination, until no more shedders
* Day 0 for Group 2 will occur after Group 1 has completed the study and the facility has been decontaminated ** The first stool of each day will be collected continuously until all subjects cease shedding (3 consecutive PCR-negative samples), or until all subjects in their respective group have at least 28 days of contained follow-up, whichever comes first.				

4 Analysis Populations

4.1 Intention -to -Treat Population

The Intention-to Treat (ITT) population is defined as all subjects who successfully complete the screening visit and are invited to participate in the study.

4.2 Total Vaccinated Population

The Total Vaccinated population (TVP) is defined as all subjects who are in the ITT population and who received study vaccine. Drop out from ITT to TVP will be described.

4.2.1 Total Vaccinated Population – negative for poliovirus she dding at baseline

In the event that any subject returns a pre-vaccination stool sample that is positive for poliovirus, an additional study population (TVPn) will be defined as all subjects in the TVP who were negative for poliovirus shedding at all assessments prior to vaccination.

4.3 Per -Protocol Population

The Per-Protocol (PP) population consists of all eligible study participants who are in the TVP and excludes those subjects who meet either of the following criteria:

- Any disease or therapy that could significantly affect the subject's immune status.
- Administration of any vaccine other than the study vaccine within 28 days of receipt of study vaccine and during the entire study period. Subjects receiving any other vaccine prior to the day 28 time point will be removed from the per-protocol population.

All deviations and violations occurring in the study will be reviewed prior to database lock and classified as either minor or major.

The TVP will be used for primary safety and shedding analysis and the PP population for immunogenicity analysis; all immunogenicity analyses (primary and secondary) will be repeated in the TVP. Demographics will be shown for all three populations. Immunogenicity and shedding analyses will be repeated in the TVPn population, if it differs in content from the Total Vaccinated population. In the event that any of the study populations contains exactly the same subjects as another study population, summaries will be prepared for only one population, and a footnote will indicate the populations that are equal to one another.

5 Statistical Considerations

5.1 General Principles

In this study, all analyses will be descriptive. No formal hypothesis will be tested. Unless otherwise specified in this document, a two-sided type I error rate of alpha = 0.05 will be used for inferential methods, such as for the generation of confidence intervals.

Unless otherwise specified, descriptive statistics (n, mean, median, standard deviation, minimum, maximum) will be used to describe continuous variables, and frequencies and percentages will be used to describe categorical variables.

Statistical analyses will be generated using SAS®, version 9.4 or above, or R version 3.3.0 or greater for specific analyses. An appendix to the statistical report will list all R packages and versions used in the report generation.

Unless otherwise specified, tables will include presentation of results by time point and by group, and overall, when possible. Figures and tables which provide more than summary statistics may include neutral explanatory text to aid in their description of the underlying data.

Unless otherwise noted, "baseline" will refer to the day of vaccination for all participants in all groups.

5.2 Randomization

Thirty (30) subjects will be recruited to this study (15 subjects in each of two sequential stages), and randomized to Group 1 or Group 2 in a 1:1 ratio. To avoid risk of transmission between subjects receiving different candidate vaccines, the study will be conducted without *individual* randomization and with each candidate vaccine sequentially. The first 15 subjects will all be enrolled in the same Group and receive the same nOPV2 candidate and the next 15 subjects will be enrolled in the other Group and receive the corresponding nOPV2 candidate; all site staff will be blinded to which candidate is administered to which cohort. Candidate 1 or 2 will be assigned to the first group randomly.

Study staff will be blinded to the vaccine candidate administered, and will be blinded for individual shedding results of the participants of a group until end of containment of this group. At that time point study staff will be notified of any remaining shedding to be able to provide these subjects with adequate instructions to take home.

CDC personnel will be responsible for communicating shedding results (any positive vs. all negative) to site personnel as stool samples are received, processed, and assayed. Subjects will be released from containment when all subjects within the group have at least 3 consecutive stool samples which are PCR-negative for poliovirus shedding, or when 28 days of post-vaccination follow-up have been achieved for all subjects, whichever comes first.

5.3 Missing Data

Data values that are identified by quality control procedures to be spurious will not be used in final analyses of trial data. Spurious data which have been removed from analysis will be clearly documented, with justification provided.

For primary and secondary objectives, imputation will not be used.

6 Planned Analyses

6.1 Subject Disposition

A CONSORT diagram will be used to describe number of subjects: enrolled by study group, receiving immunization, with complete ascertainment of serology endpoints, and included in the three primary analysis cohorts (ITT, TVP, PP).

This diagram will be supplemented with a table summarizing the number and percentage of those not receiving study vaccination (and hence replaced), those who withdrew from the study, and the reasons for withdrawal.

6.2 Demographic, Baseline Characteristics, and Concomitant Medications/Vaccinations

Descriptive statistics will be provided per group for demographic characteristics (age, height, weight, race, and gender) for each study population. All other initial subject characteristics (e.g., clinical laboratory values, physical examination, medical history, concomitant medications) will be provided by group for participants in the Total Vaccinated Population.

Baseline vital signs (heart rate, systolic/diastolic blood pressure, body temperature) will be summarized as continuous variables. The frequency of clinically relevant vital signs will be summarized by group and time point.

Summary statistics for clinical laboratory values will be reported by group and time point for baseline chemistry, hematology and coagulation, separately for abnormal and normal measurements, and will include a categorical summary of clinically relevant out-of-range values. Concomitant medications will be coded and summarized using the WHO_DRUG Dictionary as well as listed, separately for prior (not ongoing at study start) and concomitant (ongoing at study start or started while on study) medications.

Prior polio vaccination history will be tabulated. Any non-study vaccine given during the study period will be listed.

Medical history will be summarized and listed.

Physical examination findings at baseline and performance of the psychological examination will be summarized by group.

Baseline immunogenicity assessments, including seroprotection rates and type 2 $\log_2$ neutralizing antibody titers, will be summarized as described in Section 6.7.

6.3 Protocol Deviations

If protocol deviations occur, a listing of protocol deviations will be prepared, including classification of status as either major or minor.

6.4 Vaccine Administration and Sample Collection

A summary of vaccine administration will be prepared, describing the percent of subjects enrolled that received vaccination, and the time (days) between screening and vaccination visits.

A summary table and listing will be prepared to describe the percent of TVP subjects providing evaluable (able to be evaluated via PCR for detection of poliovirus) stool samples for screening and each post-vaccination day of the study.

6.5 Safety Analyses

The TV population will be used for all safety summaries. Analyses described below will be performed for solicited and unsolicited AEs as well as for SAEs and severe AEs.

The original terms used in the designated sections of the eCRFs by Investigators to identify AEs will be fully described and coded according to the Medical Dictionary for Regulatory Activities (MedDRA), including System Organ Class (SOC) and Preferred Term (PT). The following categories will be used by the Investigator to describe the causality assessment:

> **Unrelated** – there is not a reasonable possibility that the study vaccine caused the AE.
>
> **Unlikely** – suggests that only a remote connection exists between the study vaccine and the event. Other conditions, including concurrent illness, progression or expression of the disease state or reaction to concomitant medication, appear to explain the AE.
>
> **Possible** – suggests that the association of the AE with the study vaccine is unknown, however the event is not reasonably supported by other conditions.
>
> **Probable** – suggests that a reasonable temporal sequence of the AE with vaccine administration exists and, in the Investigator's clinical judgment, it is likely that a causal relationship exists between the vaccine administration and the AE, and other conditions (concurrent illness, progression or expression of the disease state, or concomitant medication reactions) do not appear to explain the AE.

All AEs will be summarized by type, seriousness, severity, causality, by group and overall. AEs will be summarized with tables at the subject level, where a subject contributes to the total once under the maximum severity/relationship of the event type. A summary table of total number of AEs by SOC, PT, severity, and causality will be prepared as well. An adverse event summary table will be prepared summarizing the number of subjects with an AE, an SAE (by type), any AE by severity, any serious or severe AE, and AEs leading to study/treatment withdrawal. All adverse events will be presented in a listing, and separate listings will be developed for AE categories of special interest, as described below.

6.5.1 Primary Safety Endpoi nts

- SAEs and severe AEs will be summarized by group, by causality, severity, and reasons for SAE classification (wherever such subcategories are relevant). These events will also be listed.
- AE summary tables will be prepared for those subjects who died or withdrew from the study due to an AE.

6.5.2 Secondary Safety Endpoints

- Solicited AEs, collected within 7 days following vaccination, will be reported by group, by type, causality, severity, and by term.
- Unsolicited AEs will be reported group, by suspected causality and severity, and by SOC and PT.
- A listing of AEs including verbatim term, start/stop date, coded terms, severity, causality, actions taken, outcomes, and type (solicited vs. unsolicited) will be provided.
- AEs (solicited and unsolicited) SOC/PT by group, and those causally related to study vaccine will be provided as supplementary tables.

Abnormal chemistry, hematology, and coagulation values at each time point will be summarized and displayed graphically with boxplots. Summary tables will include the continuous raw values, the change from baseline, and the proportion of subjects with clinically relevant abnormalities for each group and post-vaccination day. Boxplots will be shown for the raw value by lab parameter and study collection day, and will include separate symbols to depict values out of range and those out of range and clinically relevant. A listing of deviations from normal laboratory ranges, including clinical relevance, will also be prepared. Clinical laboratory test values will be evaluated according to Common Terminology Criteria for Adverse Events (CTCAE) version 4.03 (toxicity grades) or in accordance with the normal ranges of the clinical laboratory (below, within, or above normal range) for parameters for which no toxicity grades are defined. Pregnancy test results will also be listed. The limits of quantification will be substituted for the actual value, where necessary.

6.5.3 Additional Safety Endpoints

Clinically relevant vital sign (heart rate, systolic/diastolic blood pressure, body temperature) abnormalities will be summarized by group and by time point (day, prior to/following vaccination). The percentage of subjects in each group at each time point with clinically relevant values beyond clinically important limits will be summarized. Vital signs will be summarized by time point, including change from baseline, by group.

Physical exam results throughout the study period will be summarized by presence/absence of abnormal findings, and abnormal physical exam findings will be listed.

For all safety summaries and listings, any subject receiving a non-study vaccination will have safety data following receipt of the non-study vaccine removed from TV population analyses. For each safety category (AEs, safety labs, vital signs and physical exam findings), separate listings will be created for safety data from TV population members following receipt of non-study vaccine. If any listing contains data from more than 5 subjects, a table will be created to summarize the data.

6.6 Viral Shedding Analyses

All viral shedding analyses will be descriptive in nature, and no specific hypothesis is intended to be tested. The total vaccinated population will be used for viral shedding analyses. Although it is not expected, if any subject is found to be shedding poliovirus prior to vaccination, separate analyses will be conducted that include (TVP) and exclude such subjects (TVPn).

6.6.1 Primary Viral Shedding Endpoints

The number and percent of stool samples available for analysis will be summarized by study day and by group, and a listing will be produced displaying sample availability.

Categorical summaries of type 2 viral shedding positivity (positive via qPCR) and continuous summaries of viral titers ($\log_{10}$ $CCID_{50}$/g among both shedders and shedders/non-shedders combined) will be produced by group and by post-vaccination day, and these data will be listed. Each of these will be augmented with two-sided 95% confidence intervals. The Wilson score method will be used for shedding positivity proportion, and the bootstrap method (10,000 replicates, quantile method) will be used for the median viral titers.

A summary of time-to-shedding-cessation of type 2 will be prepared using Kaplan-Meier methods. The day of cessation of shedding will be defined as the day of the first sample negative for shedding after which the following two samples are also negative. If re-infection is evident in the pattern of shedding due to possible viral transmission within the contained facility, additional exploratory analyses may be conducted to assess this endpoint using only primary infections due to initial vaccination. Transmission may be evident if subjects cease shedding (3 consecutive negative samples), but then begin shedding again at a later date. Summaries will be produced using Kaplan-Meier methods, and will include quartiles and corresponding 95% confidence intervals, as well as the estimated shedding cessation rate at study days 3, 5, 7, 10, 14, 21, and 28 with corresponding 95% confidence intervals. Subjects who are positive for type 2 viral shedding at their last available assessment date with an evaluable stool sample may be right-censored. The time-to-shedding-cessation curves will be displayed graphically.

Additionally, for each subject, a viral shedding index endpoint (SIE) will be calculated as the area under the curve (using the linear trapezoidal method) of $\log_{10}$-transformed values of viral concentration in stool samples as determined using quantitative PCR (viral identity) and $CCID_{50}$ (titer) from select stool samples taken following the vaccine dose, and this continuous variable will be summarized. The stool samples used for calculation of the primary shedding index will be from Days 0 through 7 post-vaccination (with day of vaccination defined as Day 0). In the event that

subjects are still under containment beyond Day 7, they will still be providing daily stool samples. In this case, an additional AUC will be computed from Day 0 through the last sample (last day of containment). If subjects remain under containment until Day 28 post-vaccination (≥1 subjects still shedding by Day 28), then an additional AUC will be computed using only Days 0, 7, 14, 21, and 28. If a subject is missing two consecutive shedding data points involved in a specific SIE computation due to samples not provided and/or unevaluable samples, the subject will have a missing SIE value for that SIE endpoint; in all other cases, the linear trapezoidal method will be used to compute the index from available data.

Plots of the reverse cumulative distribution of the viral shedding concentrations (by study day) and each SIE will be generated. A listing will also be produced displaying the SIE and time-to-shedding-cessation data.

6.6.2 Secondary Viral Shedding Endpoint: Neurovirulence

General Features

From each subject, one stool sample (the Exploratory Endpoint Specimen [EES], defined as the last PCR-positive stool sample with a $CCID_{50}$ above the predetermined cutoff of 4.0 $\log_{10}$ $CCID_{50}$/g of stool) will be submitted to the neurovirulence assay for this endpoint. These samples will each be assayed 3 times, with inoculation occurring on separate days. Following completion of the analyses described herein, additional time points may be submitted to the neurovirulence assay in post hoc exploratory work to aid in understanding of the longitudinal features of viral shedding from these candidates. Any subjects who do not produce a sample with sufficient viral content will not contribute a sample to be analyzed for this endpoint.

In addition, from each vaccine candidate, a minimum of 3 independent replicates of the clinical supplies (vaccine) will be submitted to the single-dose form of the neurovirulence assay. These replicates will be summarized in the same manner as those samples of shed virus obtained from subject stool samples. For model-based results, only samples from study participants will be used, except as defined in the *Comparison to Clinical Supplies* subsection below.

For details of the conduct of the neurovirulence assay, refer to [1].

Validity criteria for the back-titration of inoculum are established [1] to ensure the dose of inoculum is similar across assay iterations. In the single-dose methodology defined below, an identical nominal dose level is assumed. In the event that imbalance in the actual dose level is detected through observations documented in the summaries defined below, additional exploratory analyses should be conducted to adjust and account for this imbalance. For the multi-dose format of the study, the actual titer will be used in the model, rather than the nominal level.

Analyses

Descriptive Analyses

The number and percent of subjects providing a qualifying sample will be computed for each group. The study day from which the EES is produced will be summarized as a continuous variable, and listed.

A neurovirulence endpoint assay result (NEAR) is defined as an assay result for a given EES and dose level of inoculum such that the accompanying high- and low-dose controls and the back-titration of inoculum were all within acceptable limits. Each subject and each clinical supply will have 3 repeat NEARs at the nominal single dose level of 4.0 $\log_{10}$ $CCID_{50}$. Additional NEARs at additional dose levels will also exist if the multidose format of the assay is initiated for a given subject/sample.

All results of all assays conducted will be listed, including subject ID, vaccine group, study day of stool sample assayed, assay iteration (1-3), the back-titration of inoculum, the nominal dose level of inoculum, the total number and percent of mice evaluable and paralyzed (also separately by mouse gender), the count and percent of accompanying high-and low-dose control mouse paralysis, a flag indicating if the assay lead to a re-test, a flag indicating if the assay lead to initiation of the multi-dose format of the assay, and a flag indicating whether the sample is a NEAR. Assays of clinical supplies will also be listed similarly.

For each subject, for each NEAR, the number of mice evaluable for scoring as well as the number and percent of these mice paralyzed at each dose level tested (in the event multiple dose levels are used to inoculate groups of mice with a virus population obtained from a single subject) will be summarized (overall as well as by mouse gender), for each iteration of the assay as well as aggregated within subject across assay iteration, and within vaccine group across subjects and assay iteration. The actual titer of inoculum for these samples will be summarized as a continuous variable by subject (across assay iterations) and by group (across assay iterations and across subjects), for each nominal dose level tested.

A boxplot of paralysis proportions will be prepared for each vaccine group and nominal dose level tested, wherever >2 subjects per dose level are available. The data points used to generate this figure will be individual NEARs (no aggregation across iteration). The boxplots will include the samples from clinical supplies alongside those obtained from study subjects, using the 4.0 $\log_{10}$ $CCID_{50}$ dose level. Points will be overplotted on the boxplots.

Irregularities in mouse disposition, such as reasons for euthanization prior to paralysis scoring, will be provided in a listing. Reasons for any sample re-runs, or initiation of the multi-dose format of the assay will be provided in a listing.

Model-based Analyses

If paralysis is observed among NEARs for either group, the following analysis will be conducted. (See also the section *Alternate Method for Lack of Model Fit*, below.*)* For each vaccine Group, separately, a generalized linear mixed model (GLMM) will be fitted to the binomial count of paralyzed mice for NEARs obtained at the 4.0 $\log_{10}$ $CCID_{50}$ dose level. This model is given by

$$logit^{-1}(p_{ij}) = \beta_0 + \beta_1 I_{[sex=F]} + \delta_i \qquad (1)$$

where

- β_0 is the overall mean log-odds of paralysis
- β_1 is the difference in mean log-odds of paralysis between mouse gender
- p_{ij} is the paralysis rate for sample (subject) i, assay iteration j
- $\delta_i \sim N(0, \tau^2)$ is the subject-level random effect, intended to capture overdispersion due to between-sample variability in the neurovirulence of each virus population, variability in the precise titer of inoculum, and within-subject correlation of assay replicates

In SAS/STAT software, this model may be fitted using the PROC GLIMMIX procedure. It is preferred to use METHOD = LAPLACE in the PROC GLIMMIX statement, due to better asymptotic performance of the estimators. The SAS default method for computing degrees of freedom should be utilized for statistical tests (DDFM = BETWITHIN). SAS code to fit the model defined above is given by:

```
proc glimmix data=dat method=laplace;
  class sample;
  model x/n = sex / cl solution ddfm = betwithin;
  random intercept / subject=sample;
run;
```

where "sample" is the sample number (e.g., subject identifier providing the sample), "x" is the number of mice paralyzed, "sex" is an indicator variable for the mouse gender, and "n" is the number of inoculated mice available for analysis. The dataset "dat" should contain one row for each subject for each iteration of the assay.

Model fit results should be summarized in a table, including the intercept estimate, its standard error and 95% confidence interval (based on the *t* distribution, the default for PROC GLIMMIX); similarly for the variance component. P-values for variance components will be based on the likelihood ratio test described by Molenberghs and Verbeke [2]. Additionally, the estimated mean paralysis rate at the nominal dose level (4.0 $\log_{10}$ $CCID_{50}$), $\hat{p}$, will be obtained by inverting the logit transformation with the estimated intercept, $\widehat{\beta_0}$, and the delta method will be used to obtain its standard error, from which the 95% confidence interval will be obtained, utilizing asymptotic normality, and truncated at (0, 1) if necessary.

In the event that the multiple-dose format of the assay is employed for samples from one or more subjects, due to an excessive number of mice paralyzed (as defined in [1]):

1. The methods above will be employed, omitting the additional dose levels tested for those subject(s) *with* additional dose levels tested
2. The methods will be augmented with the analyses described in the *Statistical Methods for Multi-dose Assays* section below

Statistical Methods for Multi-dose Assays

In the event that the multiple dose format of the assay is employed for samples from one or more subjects, the GLMM described in (1) above will be augmented with a term for dose level as a continuous variable, and the NEARs will be modeled as:

$$logit^{-1}(p_{ijk}) = \beta_0 + \beta_1 D_{ijk} + \beta_2 I_{[sex=F]} + \delta_i \qquad (2)$$

where

- β_0 is the overall mean log-odds of paralysis
- β_1 is the change in log-odds of paralysis associated with a unit increase in dose level of inoculum
- β_2 is the difference in mean log-odds of paralysis between mouse gender
- p_{ijk} is the paralysis rate for sample *i* at dose level *j*, in assay iteration *k*
- D_{ijk} indicates the *actual* (as opposed to nominal) dose level for sample *i*, indexed by *j*, for assay iteration *k*
- $\delta_i \sim N(0, \tau^2)$ is the subject-level random effect, intended to capture overdispersion due to between-sample variability in the neurovirulence of each virus population and within-subject correlation of assay replicates

In SAS/STAT software, this model may be fitted using the PROC GLIMMIX procedure. It is preferred to use METHOD = LAPLACE in the PROC GLIMMIX statement, due to better asymptotic performance of the estimators. The SAS default method for computing degrees of freedom should be utilized for statistical tests (DDFM = BETWITHIN). SAS code to fit the model defined above is given by:

```
proc glimmix data=dat method=laplace;
  class sample;
  model x/n = dose sex / cl solution ddfm = betwithin;
  random intercept / subject=sample;
run;
```

If only one subject for either vaccine group is available, then the random intercept term will be omitted. If subject numbers are low (e.g., ≤3), the GLMM may fail to fit due to numerical instability regarding the variance component. In this case also, the random intercept term will be omitted, and results will be presented separately for a) all results from all assay iterations, and b) all results from only the first assay iteration.

Model fit results will be summarized in a table, including coefficient estimates and standard errors, p-values from *t*-tests of coefficients, and 95% confidence intervals for coefficients (based on the *t* distribution, the default for PROC GLIMMIX) as well as the variance component. P-values for variance components will be based on the likelihood ratio test described by Molenberghs and Verbeke [2]. Additionally, the fitted curve(s) will be plotted and accompanied by a descriptive legend, with important features such as the PD_{50} denoted in text on the plot. One curve for each subject will be shown (using estimates of the random effect terms), along with the mean curve in the case of more than 1 subject.

Additionally, if any dose level produces ≥50% of mice paralyzed, the estimated dose level corresponding to a 50% paralysis rate (PD_{50}) will be computed using inverse-prediction from the

estimated model, omitting any variance components. The delta method and an assumption of asymptotic normality will be used to compute and present the standard error of this value and an accompanying 95% confidence interval, using Normal distribution critical values.

Alternate Method for Lack of Model Fit

It is expected that candidate vaccines will produce virus populations with low neurovirulence, and therefore it is possible that very few mice are paralyzed at the fixed dose level of inoculum (4.0 $\log_{10}$ $CCID_{50}$), or at lower dose levels for either or both vaccine candidates. It is also possible that few subjects shed virus in sufficient quantity to enable the assay to be conducted. In either case, it is possible that the GLMM described above may not be able to be fitted to the data. In the event that paralysis proportions are low, the variance component will be difficult to fit. If the model is unable to be fitted using SAS default values for optimization convergence criteria, the following methods will be employed, in sequence, as backup:

1. The GLMM model will be reduced to a GLM model, by omitting the variance component term, and all methods above will be used. This method ignores the overdispersion expected by combining data from heterogeneous virus populations obtained from different subjects and by combining assay iterations with different titers of inoculum, and the intrasubject correlation across assay iterations, but will produce reliable estimates of mean paralysis rates if the contributions from each subject (sample size of mice) are relatively balanced, which is expected here. In this case, statistical inference would be affected, but the analysis is intended to be descriptive, rather than inferential.
2. In the event that 1) above fails to fit (due, for example, to *no* observation of paralysis), then all analysis will be limited to the descriptive analyses presented above equation (1)

Comparison to Clinical Supplies

From each vaccine candidate, a minimum of 3 independent replicates of the clinical supplies (vaccine) will be submitted to the single-dose form of the neurovirulence assay. In order to study the change in neurovirulence from clinical supply to shed virus, the model (1) above will be augmented upon incorporation of these clinical supply data to model these NEARs as:

$$logit^{-1}(p_{hij}) = \beta_0 + \beta_1 I_{[h=2]} + \beta_2 I_{[sex=F]} + \delta_i I_{[h=1]} + \gamma_i I_{[h=2]} \qquad (3)$$

where

- h indexes virus source (h =1 = shed virus, h = 2 = clinical supply)
- i indexes sample (subject) within levels of h
- j indexes assay iteration for a given sample (subject)
- β_0 is the overall mean log-odds of paralysis for shed virus samples
- β_1 is the difference in mean log-odds of paralysis associated with clinical supply (h=2) compared to shed virus (h=1)
- β_2 is the difference in mean log-odds of paralysis between mouse gender
- p_{hij} is the paralysis rate for virus source h, sample i, assay iteration j

- $\delta_i \sim N(0, \tau^2)$ is the subject-level random effect for shed virus, intended to capture overdispersion due to between-sample variability in the neurovirulence of each virus population, variability in the precise titer of inoculum, and within-subject correlation of assay replicates
- $\gamma_i \sim N(0, \sigma^2)$ is the subject-level random effect, intended to capture overdispersion due to between-assay variation in the precise titer of inoculum and within-subject correlation of assay replicates

Note that separate variance components are used for clinical supply vs shed virus. For the data derived from clinical supply, there is only one "group" (one source of virus), and so the variance component σ^2 only captures between-iteration assay variation. In this situation, with low paralysis rates, it is possible and likely that this component is estimated to be zero. Model fit convergence criteria should be checked carefully in this case, to ensure there are no statistical identifiability issues with the primary parameter of interest, β_1. If problems with model fit are detected, the model will be considered to have failed to fit.

In SAS/STAT software, this model may be fitted using the PROC GLIMMIX procedure. It is preferred to use METHOD = LAPLACE in the PROC GLIMMIX statement, due to better asymptotic performance of the estimators. The SAS default method for computing degrees of freedom should be utilized for statistical tests (DDFM = BETWITHIN). SAS code to fit the model defined above is given by:

```
proc glimmix data=dat method=laplace;
  class sample vsource sex;
  model x/n = vsource sex / cl solution ddfm = betwithin;
  random intercept / subject=sample type=VC group=vsource;
run;
```

where "sample" is the sample number (e.g., subject identifier providing the sample, or equivalent for clinical supply), "x" is the number of mice paralyzed, and "n" is the number of inoculated mice available for analysis, "vsource" is a binary categorical (class) variable denoting whether the virus source is shed virus (reference class, *h*=1) or the clinical supply (comparator class, *h*=2), and "sex" is a categorical variable indicating the sex of the mice. The dataset "dat" should contain one row for each subject (or clinical supply sample) for each iteration of the assay.

Model fit results should be summarized in a table, including the intercept estimate. Standard errors and 95% confidence interval (based on the *t* distribution, the default for PROC GLIMMIX) will be displayed for coefficients as well as for the variance components. P-values for variance components will be based on the likelihood ratio test described by Molenberghs and Verbeke [2]. Additionally, the estimated mean paralysis rate for each virus source *h*, $\widehat{p_h}$, will be obtained by inverting the logit transformation of the corresponding linear term from (3), and the delta method will be used to obtain the standard error, from which the 95% confidence intervals will be obtained, utilizing asymptotic normality, and truncated at (0, 1) if necessary. A two-sided 95% confidence interval for the difference in paralysis rates between virus source (clinical supply vs. shed virus, for each candidate strain) will also be computed using the delta method and asymptotic normality, and truncated at (-1,1) if necessary.

As with other model-based methods above, if the model cannot be fit, remedies involve stepwise model simplification. For this model, the first simplification will be to assume a common variance term for clinical supplies and shed virus. If that model fails to fit, the remedy will be to reduce from a GLMM to a GLM via removal of both variance components. If that model cannot be fit, results will be limited to presentation of summaries described above equation (1), and augmented with a computation of the two-sided 95% confidence interval for the difference in paralysis rates between virus source using the two-sample score-based method by aggregating all samples within virus source (across subject, across assay iteration, where relevant).

6.6.3 Exploratory Viral Shedding Endpoint: Nasopharyngeal Swabs

Nasopharyngeal swabs, taken from each participant on study days 0, 3, 7, and last day of containment. Sample availability will be summarized and listed. Samples will be evaluated via rtPCR for positivity for poliovirus, and this binary variable (positive/negative) will be summarized by group and by time point, and listed.

6.7 Immunogenicity Analyses

All immunogenicity analyses will be performed with the PP population and repeated with the TV population.

6.7.1 Secondary and Exploratory Immunogenicity Endpoints

For both days 0 and 28, neutralizing antibody titers (NAbs) to polio serotype 2 will be summarized and displayed graphically (reverse cumulative distribution) by time point, using the limits of quantification as numeric values, where relevant. Continuous-variable summary statistics will be computed and presented in tabular form for NAbs, and augmented with GMT of antibody with accompanying 95% confidence intervals for both the GMT and median $\log_2$ titers. Confidence intervals for the median will be obtained using the percentile bootstrap method (n=10,000 replicates), and confidence intervals for the GMT will be obtained via normal-distribution methods.

Seroprotection and seroconversion rates will be calculated for each time point (Day 28 only for seroconversion), with accompanying 95% confidence intervals constructed using the Clopper-Pearson exact method.

Graphical representations of antibody titers will include reverse cumulative distribution plots by time point and group.

6.8 Exploratory Endpoints

Deep Sequencing

A separate analysis document will describe the exploratory analyses to be conducted on the sequence of shed virus obtained from one or more stool samples from all volunteers.

Focus will be on retention of attenuating modifications, as well as the potential relationship of sequence change to changes in neurovirulence from vaccine virus to shed virus.

7 References

[1] Novel Oral Polio Vaccine Product Development: Testing of clinical trial samples by modified mouse neurovirulence test. Viroclinics SOP VC-M140.

[2] Verbeke, G. and Molenberghs, G. (2003), "The Use of Score Tests for Inference on Variance Components," Biometrics, 59, 254–262.