Statistical Analysis Plan

Date: 10 May 2019

Author(s): Chris Gast, PhD

Protocol Number: UAM4

Protocol Title: A Phase 2, partial blind, randomized, placebo-controlled, multicenter study to evaluate the safety and immunogenicity of two novel live attenuated serotype 2 oral poliovirus vaccines candidates, in healthy adults previously vaccinated with oral polio vaccine (OPV) or inactivated polio vaccine (IPV), compared with historical controls given Sabin OPV2 or placebo.

APPROVAL SIGNATURES

Protocol Number:	UAM4
SAP Version:	1.0
Version Date:	10 May 2019

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

Christopher Gast, Ph.D. — Date: 2019/05/10

Biostatistician

Prof. Dr. Pierre Van Damme, Ph.D. — Date: 2019/05/12

Principal Investigator

Change / Release Log

Date	Version	Changes
10 MAY 2019	1.0	Initial release

Table of Contents

1 List of Abbreviations and Definitions of Terms

1.1 Abbreviations

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

2 Introduction

This document outlines the statistical methods to be implemented for the analysis of the data resulting from Protocol UAM4, *A Phase 2, partial blind, randomized, placebo-controlled, multicenter study to evaluate the safety and immunogenicity of two novel live attenuated serotype 2 oral poliovirus vaccines candidates, in healthy adults previously vaccinated with oral polio vaccine (OPV) or inactivated polio vaccine (IPV), compared with historical controls given Sabin OPV2 or placebo*. 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 2 study of the safety and immunogenicity of two novel attenuated serotype 2 oral poliovirus vaccines given to adults with either an inactivated polio vaccine (IPV) background, or a live oral polio vaccine (OPV) background. Its purpose is to provide safety, immunogenicity, and stool viral shedding data that will support continued clinical development in younger age groups, targeted at 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 design

This is a multicenter, partial blind, placebo-controlled, randomized study in 200 healthy OPV-vaccinated adults and in 48 to 132 healthy IPV-only vaccinated adults (age range 18 to 50 years), as follows:

- 50 OPV-vaccinated adults to receive 1 dose of nOPV2 candidate 1 (Group 1);
- 50 OPV-vaccinated adults to receive 2 doses of nOPV2 candidate 1 (Group 2), administered 28 days apart;
- 50 OPV-vaccinated adults to receive 1 dose of nOPV2 candidate 2 (Group 3);

- 50 OPV-vaccinated adults to receive 2 doses of nOPV2 candidate 2 (Group 4), administered 28 days apart;
- 16 to 44 IPV-only vaccinated adults to receive 2 doses of nOPV2 candidate 1 (Group 5), administered 28 days apart;
- 16 to 44 IPV-only vaccinated adults to receive 2 doses of nOPV2 candidate 2 (Group 6), administered 28 days apart;
- 16 to 44 IPV-only vaccinated adults to receive placebo (Group 7), administered 28 days apart.

3.3 Schema

Error! Reference source not found. provides a summary of the study groups and vaccines administered, including background vaccination, number of doses, and sample size.

Study code/ Country	Group N°	Age (yr)	Immunogenic Background	Candidate	Study Cohort Size	No. of Doses	Dose level
UAM4 Belgium	Group 1	18-50	OPV	nOPV2 candidate 1	50	1	~10^6
UAM4 Belgium	Group 2	18-50	OPV	nOPV2 candidate 1	50	2	~10^6
UAM4 Belgium	Group 3	18-50	OPV	nOPV2 candidate 2	50	1	~10^6
UAM4 Belgium	Group 4	18-50	OPV	nOPV2 candidate 2	50	2	~10^6
UAM4 Belgium	Group 5	18-50	IPV	nOPV2 candidate 1	16-44	2	~10^6
UAM4 Belgium	Group 6	18-50	IPV	nOPV2 candidate 2	16-44	2	~10^6
UAM4 Belgium	Group 7	18-50	IPV	Placebo	16-44	2	Sirupus simplex

Table 1. Study Schema

3.4 Study objectives

3.4.1 Primary objective

The primary objectives of the study are

- To assess the safety (serious adverse events [SAEs] and severe adverse events [AEs]) of novel monovalent live attenuated oral serotype 2 poliovirus vaccine (nOPV2) candidate 1 and novel monovalent live attenuated oral serotype 2 poliovirus vaccine (nOPV2) candidate 2 in healthy OPV-vaccinated adults, relative to historical controls given Sabin OPV2 (study UAM1);
- To compare the immunogenicity (seroprotection rate) of novel monovalent live attenuated oral serotype 2 poliovirus vaccine (nOPV2) candidate 1 and novel monovalent live attenuated oral serotype 2 poliovirus vaccine (nOPV2) candidate 2 in healthy OPV-vaccinated adults to historical controls given Sabin OPV2 (study UAM1);
    - **Seroprotection** is defined as type 2-specific neutralizing antibody titers ≥1:8 and seroprotection rate as the percentage of seroprotected subjects per group.
- To assess the safety (serious adverse events [SAEs] and severe adverse events [AEs]) of nOPV2 candidate 1 and nOPV2 candidate 2 in healthy IPV-only vaccinated adults, compared with placebo.

3.4.2 Secondary objectives

Secondary objectives are

- To assess the safety (any solicited and unsolicited AEs, laboratory assessments) of nOPV2 candidate 1 and nOPV2 candidate 2 in healthy OPV-vaccinated adults, compared with historical controls given Sabin OPV2 (study UAM1);
- To assess the safety (any solicited and unsolicited AEs, laboratory assessments) of nOPV2 candidate 1 and nOPV2 candidate 2 in healthy IPV-only vaccinated adults, compared with placebo;
- To compare the immunogenicity (seroconversion rate, median antibody titer (post-vaccination)) of nOPV2 candidate 1 and nOPV2 candidate 2 in healthy OPV-vaccinated adults, with historical control of Sabin OPV2 (study UAM1);
    - **Seroconversion** is defined as a change from seronegative to seropositive and neutralizing antibody titers of ≥1:8, and in seropositive subjects, as an antibody titer increase of ≥ 4 fold over baseline titers.
- To assess the immunogenicity (seroprotection rate, seroconversion rate, median post-vaccination antibody titer) of nOPV2 candidate 1 and nOPV2 candidate 2 in healthy IPV-only vaccinated adults.

3.4.3 Exploratory objectives

Exploratory objectives are

- to compare immunogenicity (geometric mean titer [GMT]) of nOPV2 candidate 1 and nOPV2 candidate 2 in healthy OPV-vaccinated adults, with historical control of Sabin OPV2 (study UAM1);
- to assess immunogenicity (geometric mean titer [GMT]) of nOPV2 candidate 1 and nOPV2 candidate 2 in healthy IPV-only vaccinated adults;
- to compare viral shedding following nOPV2 candidate 1 or nOPV2 candidate 2 administration in a pre-specified subset of stool samples in healthy OPV-vaccinated adults, with historical control of Sabin OPV2 (study UAM1);
- to assess viral shedding following nOPV2 candidate 1 or nOPV2 candidate 2 administration in a pre-specified subset of stool samples in healthy IPV-only vaccinated adults;

The shedding analyses may initially be conducted using a subset of samples in the 10 days following administration of each dose (e.g. day 0, 3, 5, 7, and 10) and all samples at day 14 and thereafter. The remaining samples will be stored pending evaluation of the initial analyses; PCR and $CCID_{50}$ will be conducted on these stored samples if deemed important for the complete interpretation of the study results.

Exploratory objectives will also include comparison of neurovirulence (as measured in animal model(s)) and assessment of the genetic stability via analysis of genetic sequence, including but not limited to the modified regions of shed virus in a subset of stool samples of all OPV-vaccinated adults, relative to historical control of Sabin OPV2.

Exploratory objectives will also include assessment of neurovirulence (as measured in animal model(s)) and the genetic stability via analysis of genetic sequence, including but not limited to the modified regions of shed virus in a subset of stool samples of all IPV-only vaccinated adults given doses of nOPV2 candidates 1 and 2.

3.5 Study endpoints

3.5.1 Primary endpoints

Safety: The following endpoints will be evaluated by group and overall:

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

Immunogenicity:

- seroprotection rate of type 2 polio antibodies at Day 28, following a single dose of nOPV2 candidate 1 (combined Groups 1 and 2).
- seroprotection rate of type 2 polio antibodies at Day 28, following a single dose of nOPV2 candidate 2 (combined Groups 3 and 4).

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

3.5.2 Secondary endpoints

Safety: The following endpoints will be evaluated by group and overall:

- incidence, type, causality and severity of solicited adverse events for days 0-7 in Groups 1 and 2 combined and days 28-35 in Group 2;
- incidence, type, causality and severity of solicited adverse events for days 0-7 in Groups 3 and 4 combined and days 28-35 in Group 4;
- incidence, type, causality and severity of solicited adverse events for days 0-7 and days 28-35 in Groups 5, 6 and 7;
- incidence, type, causality and severity of unsolicited adverse events throughout the study period in all groups;
- incidence, causality and description of deviations from normal safety labs at Day 0, Day 7, Day 14 and Day 28 for Groups 1 through 4, and at Day 35, Day 42 and Day 56 for Groups 2 and 4;
- incidence, causality and description of deviations from normal safety labs at Day 0, Day 7, Day 14, Day 28, Day 35, Day 42 and Day 56 for Groups 5, 6, and 7.

Immunogenicity:

- Median titers of type 2 polio antibodies at Day 28 in Groups 1 and 2 combined;
- Median titers of type 2 polio antibodies at Day 28 in Groups 3 and 4 combined;
- Median titers of type 2 polio antibodies at Day 28 in Groups 5, 6 and 7.
- Seroprotection rate and median titers of type 2 polio antibodies at Day 56 in Groups 2 and 4;
- Seroprotection rate of type 2 polio antibodies at Day 28 and at Day 56 in Groups 5, 6 and 7.
- Seroconversion rate of type 2 polio antibodies at Day 28 for Groups 1 and 2 combined and at Day 56 in Group 2;
- Seroconversion rate of type 2 polio antibodies at Day 28 for Groups 3 and 4 combined and at Day 56 in Group 4;

- Seroconversion rate of type 2 polio antibodies at Day 28 and at Day 56 for Groups 5, 6 and 7.

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

3.5.3 Exploratory endpoints

- GMT of type 2 polio antibodies at Day 28 in Groups 1 and 2 combined and at Day 56 in Group 2;
- GMT of type 2 polio antibodies at Day 28 in Groups 3 and 4 combined and at Day 56 in Group 4;
- GMT of type 2 polio antibodies at Day 28 and at Day 56 in Groups 5, 6 and 7.
- Viral shedding positivity rate (as determined using quantitative PCR) will be assessed at a pre-specified subset of the stool collection time points
- Median 50% cell culture infective dose ($CCID_{50}$; titer) of shed virus after viral extraction from PCR-positive stool samples will be assessed at a pre-specified subset of the stool collection time points
- Time-to-cessation of viral shedding will be assessed following each dose
- The extent of shedding, including quantity and duration, will be assessed among pre-specified stool collection time points following each dose.

Exploratory endpoints will include assessment for neurovirulence of shed virus (as measured in animal model(s)) and will also include assessment of the genetic stability of the modified regions of shed virus in a subset of stool samples.

3.6 Randomization

All OPV-vaccinated subjects are to receive one of the nOPV2 candidates in a single blind manner and all IPV- vaccinated subjects are to receive one of the nOPV2 candidates or placebo in a double-blinded manner.

For the whole study duration all subjects and blinded study staff responsible for safety evaluation of IPV-subjects will not have any information of what has been administered. As the placebo can be distinguished from the vaccine candidates in packaging and color, reception of the vaccines, dose preparation and administration is done by a team of unblinded study personnel. Appropriate measures are taken at the site to ensure blinding of subjects and blinded team for the whole duration of the study.

Groups are enrolled sequentially such that subjects receiving candidate 2 are enrolled first (IPV and OPV groups simultaneously), followed by subjects receiving candidate 1. IPV-only vaccinated adults are randomized 2:1 to candidate 2 (Group 6) or placebo (Group 7), respectively, until 16 to 44 Group 6 subjects are enrolled, after which randomization 2:1 to candidate 1 or placebo commences for Groups 5 and 7. For OPV-vaccinated subjects, randomization to Groups 1 and 2 to receive candidate 1 ensues following complete randomized enrollment of subjects to Groups 3 and 4. In all cases, block randomization is used to ensure balanced randomization across time.

Allocation of each subject to a given group is described in a computer-generated randomization schedule prepared prior to start of the study by Assign Data Management and Biostatistics GmbH, Stadlweg 23, 6020 Innsbruck using nQuery Advisor®. Assign also provides the site with randomization and emergency envelopes.

3.7 Data Monitoring

An independent Data and Safety Monitoring Board (DSMB) is charged with monitoring the benefit-risk and data integrity of this trial, to inform the sponsor about recommendations for future clinical trial phases.

The DSMB has established stopping rules for safety prior to study start, which are continuously assessed. The DSMB will also monitor study enrollment, particularly for IPV-vaccinated subjects, and recommend truncation and/or closure of study groups if enrollment stagnates and when current enrollment is considered sufficient to meet study objectives and no safety signals occurred. The minimum number of IPV-vaccinated subjects agreed on by DSMB per candidate vaccine is 24. With a randomization of 2:1 for placebo the minimum study cohort size for Groups 5, 6 and 7 will be 16. In case of safety signals the DSMB reserves the right to reverse the truncated enrollment. The structure and content of the DSMB reports will agreed between the sponsor and the DSMB, independent of this analysis.

4 Analysis Populations

4.1 Intention-to-Treat Population

The Intention-to Treat (ITT) population is defined as all subjects who are randomized according to randomized treatment assignment.

4.2 Total Vaccinated Population

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

4.3 Per-Protocol Population

The Per-Protocol (PP) population consists of all eligible study participants who are in the TVP and who receive all of the immunizations scheduled for the group to which they are allocated and excludes those subjects who meet any of the criteria outlined in protocol Section 5.3. All deviations and violations occurring in the study will be reviewed prior to database lock and classified as either minor or major. The PP population requires adherence to all vaccinations (including timing) for those from Groups 1, 3, 5 and 6, but subjects will not be removed from the population for missing outcome assessments; these will be handled case by case as missing data. In case of missing second vaccination of a subject in a 2-dose Group (Groups 2 and 4) but adherence to the protocol until that timepoint, all previous collected data will still be analyzed in the PP population. A subject matching items under protocol Section 5.3 will be removed from the PP analysis.

The TVP will be used for primary safety analysis and the PP population for primary immunogenicity analysis; all immunogenicity analyses (primary and secondary) will be repeated in the TVP.

5 Statistical Considerations

5.1 General Principles

The study UAM1, conducted in a pre-OPV2 cessation period (2015-2016), was specifically designed to provide active control comparator data for both safety and immunogenicity for the OPV-vaccinated adults given the novel vaccine candidates in this study. For all endpoints for OPV-vaccinated subjects, summaries of data from this study will be presented alongside corresponding summaries from UAM1. Summaries of data from IPV-vaccinated subjects will be presented according to group (each candidate, and placebo), and have no historical control companion groups. Summaries of demographics, baseline characteristics, and immunogenicity data from this study will be prepared by site and overall, to assess comparability of sites. For OPV-vaccinated subjects, comparative analyses will be limited to comparison of each vaccine to the UAM1 Sabin 2 control; for IPV-vaccinated subjects, each candidate vaccine group will be compared to the placebo group for safety endpoints.

All data will be summarized and/or listed. 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.

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 first vaccination for all participants in all groups.

To the extent possible, unintended differences in terms collected on CRFs between the two studies will be aligned for summary and presentation, with any recoding of UAM1 terms clearly specified in the study report. Coded terms will use a common coding dictionary across the two studies.

5.2 Timing of Analysis

A final analysis of all data will be performed when all data are collected and the database is locked, following a blinded (for IPV-vaccinated subjects) review of criteria for membership in the per-protocol population. The same principles for deviation classification applied to subjects from study UAM1, documented in the data review meeting report and common between UAM1 and UAM4 protocols, will be used here.

Within background vaccination cohort (OPV vs IPV), enrollment in this study was sequential, with respect to candidate. A pre-scheduled unblinded DSMB safety review of data from OPV-vaccinated (n=100) and IPV-vaccinated subjects in Group 3, 4, 6, and 7 (receiving candidate 2) was held following accumulation of 14 days post-first-dose safety data on 24 IPV-vaccinated subjects (randomized 2:1, candidate 2:placebo), to enable age de-escalation in a subsequent study in children and infants. A similar unblinded DSMB safety review was held with updated data to include 14 day post-first-dose safety data from 26 additional subjects receiving candidate 1 or placebo (Groups 1, 2, 5, and 7), to enable age de-escalation in the subsequent phase II study with candidate 1.

5.3 Missing Data

In spite of best efforts to collect complete data for all study subjects, some data will be missing at the end of the trial. For this study, data will be assumed to be missing completely at random, and

no imputation of results will occur. If evidence suggests data are not missing completely at random, additional sensitivity analyses, such as multiple imputation, may be justified and performed as additional exploratory analyses to augment the primary results.

6 Planned Analyses

6.1 Subject Disposition

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

This diagram will be supplemented with a table summarizing these same data, by site and overall, including the reasons for withdrawal or discontinuation.

A listing and summary table of reasons for screen failures (inclusion/exclusion criteria violated) will be produced.

6.2 Demographic, Baseline Characteristics, and Concomitant Medications/ Vaccinations

Descriptive statistics will be provided per group, overall and by-site, for demographic characteristics (age, height, weight, race, and gender) for each study population. All other initial subject characteristics (e.g., physical examination abnormalities, medical history, prior/ongoing medications) will be provided by group for participants in the Total Vaccinated Population.

Concomitant medications will be coded and summarized using the WHO Drug Dictionary, separately for all prior or ongoing medications as well as only those ongoing at study start, as well as listed.

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.

Baseline immunogenicity assessments, including seroprotection rates and type 2 $\log_2$ neutralizing antibody titers, will be summarized as described in Section 6.6.

Deviations from normal of baseline vital signs, physical exam, and clinical laboratory assessments will be summarized as described in Section 6.5.

6.3 Protocol Deviations

If protocol deviations occur, a listing of protocol deviations will be prepared, separately for non-subject-specific deviations and subject-specific deviations, 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 in the ITT population that received each vaccination, all vaccinations, and the time (days) between 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 each planned collection day.

6.5 Safety Analyses

The TVP will be used for all safety summaries, with data summarized according to the vaccine received. Summaries will be presented by group and by time period, as well as across time periods within group and within vaccine group (e.g., combined Groups 1 & 2); for summaries presented by post-dose time period, these will be combined within Groups 1 & 2 (for each of UAM4 and UAM1) and within Groups 3 & 4 (UAM4) prior to dose 2. Summaries will also be presented according to candidate received, regardless of background vaccination status (i.e. combined Groups 1 and 2 from M1, combined Groups 1, 2, and 5 and combined Groups 3, 4, and 6) alongside the placebo group (Group 7). For summaries of adverse events in the post-dose-2 period, only subjects receiving the second vaccine will be included in summaries.

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.

Only adverse events collected on or after the date of first vaccination will be included in summaries; all AEs, regardless of collection time point will be included in listings.

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.

Throughout, when the word "related" is used to refer to adverse events, it is to be understood to refer to those events considered either possible or probable.

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. An adverse event summary table will be prepared summarizing the number of subjects with any unsolicited AE (and by maximum severity and by maximum relationship), any SAE by type and by maximum relationship, any related SAE, any severe unsolicited AE by maximum relationship, any serious or severe AE, any AEs leading to study/treatment withdrawal, any grade ≥2 unsolicited AE, and any solicited AEs by maximum severity. All adverse events will be presented in a listing, and separate summaries and listings will be developed for AE categories of special interest, as described below.

All summaries of unsolicited adverse events will present the number and percentage with a qualifying event, including a two-sided 95% exact confidence interval for the proportion, as well as the total number of qualifying events. Summaries of solicited adverse events will be summarized similarly but will omit the total event count.

6.5.1 Primary Safety Endpoints

- Solicited adverse events:
    - Serious/severe solicited adverse events will be summarized by group, by post-vaccination time period as well as overall (under maximum severity), and by term by post-vaccination time period as well as overall (under maximum severity), as well as listed.
- Unsolicited adverse events:
    - Serious/severe unsolicited AEs will be summarized by group, maximum causality and by class (serious vs severe) under maximum causality, by post-vaccination

time period (prior to dose 2 vs following dose 2) as well as overall, and reasons for SAE classification (wherever such subcategories are relevant), as well as by SOC and PT among serious events only (and among those serious and related), and among severe events only (and among those severe and related). These events will also be listed.

- Listings will be prepared for those subjects who died or withdrew from the study due to an AE.

Two-sided p-values from the Fisher exact test of the number of subjects with an event will be presented for:

- Severe solicited events (overall, and by event term)
- Any serious unsolicited event
- Any serious and related unsolicited event
- Any severe unsolicited event
- Any severe and related unsolicited event

These comparisons will be conducted for the following groups:

- UAM4 Groups 1 & 2 vs UAM1 Groups 1 & 2
  - Post-dose-1 and overall, TVP
- UAM4 Group 2 vs UAM1 Group 2
  - Post-dose-2, TVP subjects who received both doses
- UAM4 Groups 3 & 4 vs UAM1 Groups 1 & 2
  - Post-dose-1 and overall, TVP
- UAM4 Group 4 vs UAM1 Group 2
  - Post-dose-2, TVP subjects who received both doses
- UAM4 Group 5 vs UAM4 Group 7
  - Post-dose-1, post-dose-2, and overall, TVP
- UAM4 Group 6 vs UAM4 Group 7
  - Post-dose-1, post-dose-2, and overall, TVP

Comparisons for both solicited and unsolicited events will be conducted separately for post-dose-1 and post-dose 2, as well as following any dose. Comparisons involving post-dose-2 data will include only subjects receiving both doses.

6.5.2 Secondary Safety Endpoints

- Solicited AEs, collected within 7 days following vaccination, will be summarized by group overall and by term, (any event regardless of severity, and by severity according to maximum severity, across terms and for each term), for each post-vaccination time period as well as overall, as well as listed.
- Unsolicited AEs will be summarized by group, by relationship under maximum causality, by severity under maximum severity, by post-vaccination time period as well as overall, as well as by SOC and PT, and by SOC and severity (maximum severity). Separate tables by SOC and PT will be produced for moderate or greater events, and for moderate or greater related events only. A table will be produced summarizing preferred terms for those terms arising in >2% of subjects, sorted in descending order.
- A listing of AEs including verbatim term, start/stop date, timing (post-dose 1 vs post-dose 2), including post-prior-vaccination onset day, coded terms, severity, causality, actions taken, outcomes, and type (solicited vs. unsolicited) will be provided.

Clinical laboratory test values will be evaluated according to Common Terminology Criteria for Adverse Events (CTCAE) version 4.03 (toxicity grades using ranges for laboratory abnormalities) 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. The limits of quantitation will be substituted for the actual value, where necessary.

Clinical laboratory abnormalities will be summarized by lab parameter, time point and grade (with separate categories for high and for low values of a given lab value, where relevant), as well as in a shift table displaying the frequency of post-baseline abnormalities by grade for each time point cross-tabulated with the classification at baseline, for each time point for each lab parameter, maximum grade across lab parameter within each time point, and maximum grade across post-baseline time points and lab parameter. A separate table will summarize according to grade only those values considered to be clinically relevant. Additional summary tables will be produced including the continuous values and the change from baseline for each group and post-vaccination day. Boxplots will display 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, including grade and clinical relevance, will also be prepared.

While all safety lab abnormalities will be summarized, for each of the specific safety lab parameters listed below, the rate of moderate or greater post-baseline abnormalities will be presented at the subject level overall as well as by timepoint, and by abnormality grade (using the maximum grade post-baseline, where relevant), and accompanied by a two-sided 90% confidence interval using the Miettinen and Nurminen method for the risk difference of each vaccine candidate vs. the comparator group: ALT, AST, CK, GGT, bilirubin and albumin. For IPV-vaccinated subjects (Groups 5 and 6), comparisons will be made relative to the placebo

group (Group 7), and for OPV-vaccinated subjects (Groups 1 – 4), comparisons will be made relative to the corresponding UAM1 control (for lab values assessed in both studies), combining groups where possible, as described above (e.g., Group 1 & 2, post-dose-1).

Pregnancy test results will be listed.

Abnormal vital signs (heart rate, systolic/diastolic blood pressure, body temperature) will be summarized by group and by time point (day, prior to/following vaccination). The summary will also include the number and percent of subjects with abnormal and clinically relevant vital signs. Vital signs will be listed, and summarized as continuous variables by time point, including change from baseline, by group, and change pre- to post-vaccination for vaccination visits.

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

6.6 Immunogenicity Analyses

All immunogenicity analyses will be performed with the PP population and repeated with the TV population. Descriptive summaries will be computed overall within group, and by site within group. In all cases, data from Groups 1 & 2 (for UAM4 and UAM1) will be combined prior to dose 2, as well as data from Groups 3 and 4 (UAM4).

6.6.1 Primary, Secondary, and Exploratory Immunogenicity Endpoints

Descriptive Analysis

For all time points where neutralizing antibody titers (NAbs) to polio serotype 2 are obtained, these will be summarized and displayed graphically (reverse cumulative distribution) by time point. Continuous-variable summary statistics will be computed and presented in tabular form for $\log_2$ NAbs and augmented with the 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). Separately for each group and time point, the GMT will be computed in an intercept-only model using likelihood-based methods accounting for censoring, implemented in SAS PROC LIFEREG, using left- and right-censoring at assay LLOQ (2.5 $\log_2$) and ULOQ (10.5 $\log_2$), respectively, and a Normal error distribution assumption on the $\log_2$ scale and confidence intervals for the GMT will be obtained via reverse-transformation of the $\log_2$ mean and its corresponding two-sided 95% confidence intervals. For each post-baseline time point, the geometric mean fold rise (GMFR) will be computed via the $\log_2$ differences from baseline (and from prior visit, for two-dose recipients) with two-sided 95% confidence intervals based on the *t*-distribution, followed by the antilog transformation.

Seroprotection and seroconversion rates will be calculated for each time point (post-vaccination time points only for seroconversion), with accompanying 95% confidence intervals constructed using the Clopper-Pearson exact method. These binary endpoints will be further summarized corresponding to baseline serostatus. Seroconversion will additionally be computed among the subset of subjects where seroconversion was possible to observe (less than or equal to 4-fold from assay ULOQ at baseline).

Seroprotection rates at baseline will be accompanied by a two-sided Miettinen and Nurminen 95% confidence interval for the difference for Groups 1 & 2 combined as well as Groups 3 & 4 combined vs the UAM1 control (Group 1 & 2 combined). Similarly, comparisons of baseline seroprotection rate between Group 5 vs 7 and Group 6 vs 7 will be conducted. Comparison of baseline GMTs among these same groups will be conducted by computing the GMT ratio and corresponding two-sided 95% confidence interval obtained from modeling the $\log_2$ NAb titer as a function of group using SAS PROC LIFEREG, as described above. For IPV-vaccinated subjects, this will include a fixed effect for study site, utilizing SAS LSMEANS for pairwise comparisons. For by-site summaries, comparisons will only be available at the site in common between UAM4 and UAM1.

Comparative Analysis

The primary immunogenicity endpoint, seroprotection after a single dose of each vaccine candidate in OPV-vaccinated subjects, will be formally compared to the corresponding endpoint from UAM1 via a non-inferiority test of each of the novel candidates to the UAM1 Sabin 2 control, each using one-sided $\alpha = 0.025$ and a non-inferiority margin of -10%.

Secondary endpoints will be compared to corresponding endpoints from UAM1 in the same manner, using one-sided $\alpha = 0.025$ and considering margins of -10% for binary endpoints (seroconversion), and 2/3 for median antibody titer and GMT ratios. Non-inferiority margins for these secondary endpoints are not selected based on any specific criteria and are simply intended as benchmarks for unpowered comparisons.

In each case, non-inferiority comparisons will be conducted by computing the difference/ratio in endpoint between each candidate vaccine and the UAM1 Sabin 2 control (novel minus comparator), with corresponding two-sided $\alpha = 0.05$ confidence intervals, and comparing the lower confidence limit to the non-inferiority margin. Miettinen and Nurminen confidence intervals will be used for binary comparisons, asymptotic normal-based methods will be used for GMTs as described above using SAS PROC LIFEREG and incorporating a covariate for the baseline $\log_2$ NAb level, and bootstrap methods will be used for median titers and other continuous endpoints.

Immunogenicity among IPV-vaccinated subjects (Groups 5 through 7) will be summarized descriptively but will not be compared between groups.

The non-inferiority comparison of the primary immunogenicity endpoint (post-first-dose seroprotection rate) will be based on the combined data from each site in this study compared to the single-site data from the historical control study, UAM1, for OPV-vaccinated subjects. In order to assess any between-site differences within the current study (UAM4) and the ability to combine these in the comparison to the historical control, both seroprotection rate and GMTs will be compared between sites, within candidate vaccine arm (for OPV- and IPV-vaccinated groups):

- Both the pre-vaccination and post-vaccination seroprotection rates in UAM4 will be compared between sites by computing the two-sided 90% Miettinen and Nurminen confidence interval for the difference (corresponding to a two-sided level $\alpha = 0.10$ test). The pre-vaccination comparison will be made across all subjects in Groups 1 through 4, and also across all subjects within Groups 5 through 7. Post-vaccination comparisons will be made within vaccine candidate (within Groups 1 & 2, within Groups 3 & 4, within Group 5 and within Group 6). If 0 is not contained in the confidence interval for the post-dose-1 difference in seroprotection rate for either candidate in the OPV-vaccinated groups, comparisons for all immunogenicity endpoints will additionally be conducted among subjects only from the site in common between both studies. Additional investigations to identify the source of the differences may be conducted, e.g. subset comparison of seroprotection rates, seroconversion rates, and GMT ratios among those initially seronegative vs those initially seropositive to evaluate the role of differences in baseline immunity, and/or regression modeling.
- Between-site comparisons of GMTs will be based on a two-sided 90% confidence interval for the GMT ratio (corresponding to a two-sided level $\alpha = 0.10$ test), using SAS PROC LIFEREG to incorporate censoring at assay LLOQ and ULOQ, incorporating a covariate for baseline $\log_2$ NAb for post-baseline measurements, and using the antilog transformation for the $\log_2$ mean difference, as described above, and additionally assuming a common error variance among sites. As above, the pre-vaccination comparison will be made across all subjects in Groups 1 through 4, and separately across all subjects within Groups 5 through 7, while the post-vaccination comparisons will be made within vaccine candidate (within Groups 1 & 2, within Groups 3 & 4, within Group 5 and within Group 6). If the CI for the baseline-adjusted GMT ratio between sites post-dose-1 for either candidate among OPV-vaccinated subjects does not contain 1, the adjusted GMT ratio and corresponding CI for the GMT comparison of each candidate to the corresponding UAM1 control will additionally be conducted among subjects only from the site in common between both studies. Additional investigations to identify the source of the difference may be conducted.

6.7 Viral Shedding Analyses

Viral shedding analyses will be descriptive in nature and will include hypothesis tests contrasting nOPV2 recipients with mOPV2 recipients as defined below. The total vaccinated population will be used for viral shedding analyses following the first dose, and the per-protocol population will be used for viral shedding analyses following the 2nd dose. As above, post-dose-1 (pre-dose-2) samples will be combined within vaccine candidate and prior vaccination history, where relevant (combined Groups 1 & 2 for each of UAM1 and UAM4, combined Groups 3 & 4 for UAM4).

6.7.1 Exploratory Viral Shedding Endpoints

Throughout, viral shedding analyses will be conducted by group and dose number. The descriptive summaries below will also be computed separately by site. Between-group inferential analyses will additionally be conducted among the site in common between UAM1 and UAM4, for OPV-vaccinated subjects, if the post-dose-1 between-site GMT comparison indicates a significant difference in immunogenicity between sites. The shedding analyses will be conducted considering all available samples, collected at Days 0-10, 14, 21, 28, and 42 following each dose (with day of vaccination defined as Day 0) and additional samples collected, as necessary, to define the end of shedding for participants. All viral shedding data will be listed, including both sample collection and assay results.

Categorical summaries of type 2 viral shedding positivity (positive via PCR) 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. Each of these will be augmented with two-sided 95% confidence intervals. The Clopper-Pearson exact method will be used for shedding positivity rate, and the bootstrap method (10,000 replicates, quantile method) will be used for the median $\log_{10}$ viral titers. Subjects who are PCR-negative for viral shedding at a given time point will be assigned the value of 0 for computations involving the $\log_{10}$ $CCID_{50}$/g, and values PCR positive for type 2 poliovirus but below LLOQ (2.75 $\log_{10}$ $CCID_{50}$/g) or above ULOQ (8.25 $\log_{10}$ $CCID_{50}$/g) will be assigned the corresponding value of LLOQ or ULOQ.

For each subject, the extent of viral shedding 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 samples taken through 28 days following each vaccine dose, and this continuous variable will be summarized. If a subject is missing two consecutive shedding data points among those intended to be tested for this computation due to samples not provided and/or unevaluable samples, the subject will have a missing value. This requirement regarding consecutive missing samples may be relaxed at the analysis stage if sample size for this computation is significantly impacted. Additionally, the Shedding Index Endpoint (SIE) will be computed as the arithmetic mean of $\log_{10}$ $CCID_{50}$/g obtained from Days 7, 14, 21, and 28. Here, a single missing daily value will lead to a missing

SIE value, although windows of +/- 2 days will be applied to sample availability. If samples are available at both +1 (+2) *and* -1 (-2), then the mean $\log_{10}$ $CCID_{50}$/g from these two days will be used for this computation. This will be summarized as a continuous variable using the same methods for daily viral shedding described above.

Additionally, a summary of time-to-shedding-cessation 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. Summaries will be produced using Kaplan-Meier methods, and will include quartiles and corresponding 95% confidence intervals, as well as the estimated shedding 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, or who fail to cease shedding prior to a 2nd dose will be right-censored at the last time point they were known to be shedding virus. The time-to-shedding-cessation curves will be displayed graphically, and the comparison of the difference in time to shedding cessation curves between candidate groups in study M4 and corresponding controls subjects from study UAM1 will be conducted by computing the two-sided log-rank p-value.

A comparison of the extent of shedding among OPV-vaccinated subjects to the corresponding control data from study UAM1 will be conducted by computing the difference in viral shedding rate per study day with corresponding two-sided 95% Miettinen and Nurminen confidence intervals for the difference, and augmented with the difference in median viral shedding for the daily shedding values ($\log_{10}$ $CCID_{50}$/g) as well as the AUC and SIE alongside bootstrap-based (n=10,000 replicates each) two-sided 95% confidence intervals for the difference. A listing will include the computed time-to-shedding-cessation and the corresponding censoring indicator, the AUC, and the SIE values.

Plots of the reverse cumulative distribution of the viral shedding concentrations (by study day) and each aggregate index will be generated.

6.7.2 Exploratory Analysis of Neurovirulence of Shed Virus

Stool samples positive for type 2 viral shedding of sufficient quantity to enable an evaluation of the neurovirulence of the virus in a transgenic mouse assay may be selected from the stool samples processed through the PCR and $CCID_{50}$ assays. Not all such samples will be assessed for neurovirulence, and sub-sampling of groups or sub-sampling within groups may occur. The sub-sampling methodology used, if any, will be fully described in the clinical study report. It is anticipated that sampling will focus on Exploratory Endpoint Samples (EESs), 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, and a comparative evaluation of EES from OPV-vaccinated subjects in UAM1 and in UAM4 (both candidate vaccines in M4 vs the common M1 control) will be conducted, and the

results from IPV-vaccinated subjects will be descriptive (including non-comparative model-based methods described below).

For details of the conduct of the neurovirulence assay, refer to [1]. Validity criteria for the back-titration of inoculum are established to ensure the dose of inoculum is similar across assay iterations. In the single-dose assay statistical 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.

Samples assessed for neurovirulence in this assay will be summarized as described below.

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 are all within acceptable limits, per SOP. Each subject with a sample tested is anticipated to have 1 NEAR at the nominal single dose level of 4.0 $\log_{10}$ $CCID_{50}$/g. Additional NEARs at additional dose levels may also exist if the multidose format of the assay is initiated for a given subject/sample.

Assay results will be listed, including subject ID, vaccine group, study day of stool sample assayed, assay repeat number, 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.

For NEARs, 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), within vaccine group across subjects.

A boxplot of paralysis proportions will be prepared for each vaccine group wherever >2 subjects per dose level are available. Only NEARs using the 4.0 $\log_{10}$ $CCID_{50}$/g dose level will contribute to this summary. Points will be overlaid on the boxplots.

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 sufficient quantities of EES and, among them, sufficient mouse paralysis is observed, model-based analyses will be used to supplement the descriptive summaries described above. (See also the section *Alternate Method for Lack of Model Fit*, below.*)* It is currently known that, prior to the amplification step required to conduct the assay, only 2 post-dose-1 EES are available from the historical control subjects, UAM1. It is therefore likely that model-based analyses will fail, and/or will fail to provide additional insight beyond the descriptive summaries previously described, however, model-based analyses will still be attempted.

For each vaccine group, a generalized linear mixed model (GLMM) will be fitted to the binomial count of paralyzed mice for NEARs obtained at the 4.0 $\log_{10}$ $CCID_{50}$/g dose level. This model is given by

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

where

- β_0 is the overall mean log-odds of paralysis for male mice
- β_1 is the difference in mean log-odds of paralysis between mouse gender (females minus males)
- p_i is the paralysis rate for sample (subject) i
- $\delta_i \sim N(0, \tau^2)$ is the subject-level random effect, intended to capture overdispersion due to between-sample variability in the neurovirulence of each virus population, including variability in the precise titer of inoculum

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 sex;
  model x/n = sex / cl solution ddfm = betwithin;
  random intercept / subject=sample;
  lsmeans sex / ilink;
run;
```

where "sample" is the sample number (e.g., subject identifier providing the sample), "x" is the mouse gender-specific number of mice paralyzed, "n" is the mouse gender-specific number of

inoculated mice available for analysis, and "sex" captures the mouse gender. The dataset "dat" should contain one row for each subject, for each mouse gender. Additional options to the SAS procedure may be necessary to obtain all necessary output.

Model fit results should be summarized in a table, including coefficient estimates and standard errors, p-values from *t*-tests of coefficients, 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 per gender at the nominal dose level (4.0 $\log_{10}$ $CCID_{50}$/g), will be obtained by inverting the logit transformation, 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. A SAS LSMEANS statement will be used to obtain log-odds of paralysis averaged over mouse gender, and the corresponding probability and its 95% CI will be obtained via these same methods.

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:

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

where

- β_0 is the overall mean log-odds of paralysis for male mice
- 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 (females minus males)
- p_{ij} is the paralysis rate for sample *i* at dose level *j*
- D_{ij} indicates the *actual* (as opposed to nominal) dose level for sample *i*, indexed by *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

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 sex;
  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. Additional options to the SAS procedure may be necessary to obtain all necessary output.

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.

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 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)

Comparative Analysis for OPV-vaccinated subjects

In order to compare the neurovirulence of shed virus between each candidate vs the control, the following model will be fitted:

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

where

- h indexes virus (h =1 = control vaccine [M1], h = 2 = candidate vaccine [M4])
- i indexes sample (subject) within levels of h
- β_0 is the overall mean log-odds of paralysis for shed virus samples for male mice, control vaccine
- β_1 is the difference in mean log-odds of paralysis between the two vaccines
- β_2 is the difference in mean log-odds of paralysis between mouse gender (females minus males)
- p_{hi} is the paralysis rate for virus source h, sample i
- $\delta_i \sim N(0, \tau^2)$ 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, including variability in the precise titer of inoculum, and is assumed common among all vaccine groups (but estimated uniquely in models for candidate 1 vs control, and candidate 2 vs control)

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 / s cl solution ddfm = betwithin;
  random intercept / subject=sample;
  lsmeans vsource sex / cl ilink oddsratio e;
run;
```

where “sample” is the sample number (e.g., subject identifier providing the sample), “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 candidate vaccine or control, 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 mouse gender. Additional options to the SAS procedure may be necessary to obtain all necessary output.

Model fit results should be summarized in a table, including coefficient estimates, standard errors, p-values for *t*-tests of coefficients, 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. For each candidate vaccine, the odds ratio (paralysis in each candidate vaccine group relative to the control group from study M1) and its corresponding two-sided 95% CI and two-sided p-value will be presented. 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 from the results of the SAS LSMEANS statement, 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. The odds ratio of paralysis between mouse genders, assumed common among vaccines groups, will also be presented as described for the gender-averaged probability.

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 reduce from a GLMM to a GLM via removal of the variance component. If that model cannot be fit, results will be limited to presentation of summaries described above equation (1) and augmented with the two-sided Fisher exact p-value arising from the cross-tabulation of paralyzed mice within each group, regardless of mouse gender and subject identifier, and the p-value will be regarded as approximate. If substantial imbalance in mouse availability rate is observed, additional sensitivity analyses may be added.

6.8 Other 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 Differences from the protocol

Summaries of serum neutralizing antibody fold rise (e.g., GMFR) were not clearly specified in the protocol, and are included in the description of immunogenicity summaries to be computed, in Section 6.6.1.

8 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.