STATISTICAL ANALYSIS PLAN

Protocol Number and Title:	**AVXS-101-CL-101 (formerly AVXS-101)** **Phase I Gene Transfer Clinical Trial for Spinal Muscular Atrophy Type 1 Delivering AVXS-101**
Protocol Version and Date:	**Version 15.0 / 22 November 2016**
Author(s):	**Changsheng Wang** **Associate Director of Biostatistics**
SAP Version:	**1.0**
SAP Version Date:	**01-Dec-2016**

Revision History

Version #	Date	Document Owner	Revision Summary
1.0	22-Nov-2016	Changsheng Wang	Initial Release Version

I confirm that I have reviewed this document and agree with the content.

APPROVALS	
Changsheng Wang, Ph.D. Associate Director of Biostatistics	Date (dd-Mmm-yyyy)
Melissa Menier, Biostatistician Reviewing Biostatistician	Date (dd-Mmm-yyyy)
Doug Sproule, M.D. Vice President, Clinical Development	Date (dd-Mmm-yyyy)
Courtney Wells Director, Head of Clinical Operations	Date (dd-Mmm-yyyy)
Sukumar Nagendran, M.D. Senior Vice President, Chief Medical Officer	Date (dd-Mmm-yyyy)
James L'Italien, Ph.D. Senior Vice President, Chief Regulatory and Quality Officer	Date (dd-Mmm-yyyy)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. GLOSSARY OF ABBREVIATIONS AND DEFINITIONS

1.1. Glossary of Abbreviations

Abbreviation	Description
AE	Adverse Event
ADR	Adverse Drug Reaction
ATC	Anatomical Therapeutic Chemical
ANOVA	Analysis of Variance
AAV9	Adeno-associated virus serotype 9
ALT	Alanine Aminotransferase
AST	Aspartate Aminotransferase
BDRM	Blinded Data Review Meeting
BLQ	Below Levels of Quantification
BMI	Body Mass Index
CI	Confidence Interval
CRF/eCRF	Case Report Form/ electronic Case Report Form
CTCAE	Common Terminology Criteria for Adverse Events
CSR	Clinical Study Report
CV	Coefficient of Variation
DMC	Data Monitoring Committee
DSMB	Data Safety Monitoring Board
ECG/EKG	Electrocardiogram/Electrocardiogram
EES	Efficacy Evaluable Set
ELISAs	Enzyme-Linked Immunosorbent Assays
ELISpot	Enzyme-Linked ImmunoSpot
FAS	Full Analysis Set
FDA	U.S. Food and Drug Administration
GCP	Good Clinical Practice
GPP	Good Pharmacoepidemiology Practice
HEENT	Head, eyes, ears, nose, throat
ICH	International Conference on Harmonization
ITT	Intent-To-Treat
IVRS	Interactive Voice Response System
IWRS	Interactive Web Response System
Max	Maximum
MedDRA	Medical Dictionary for Regulatory Activities
Min	Minimum
mITT	Modified-Intent-To-Treat
N/A	Not Applicable
NA	Not Applicable
NCI	National Cancer Institute
NTF	Note-To-File
OR	Observational Research

Abbreviation	Description
PASS	Post Authorization Safety Study
PAES	Post Authorization Efficacy Study
PBMC	Peripheral Blood Mononuclear Cells
PCS	Potentially Clinically Significant
PD	Pharmacodynamic
PDS	Pharmacodynamic Set
PK	Pharmacokinetic
PKS	Pharmacokinetic Set
PPS	Per Protocol Set
PT	Preferred Term
QC	Quality Control
QTc	Corrected QT Interval
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD	Standard Deviation
SDTM	Study Data Tabulation Model
SE	Standard Error
SI	Standard International System of Units
SMN	Survival Motor Neuron
SOC	System Organ Class
SOP	Standard Operating Procedure
SS	Safety Set
TEAE	Treatment Emergent Adverse Event
TOST	Two One-Sided Tests
TLF	Table, Listing and Figure
VAS	Visual Analog Scale
WHO	World Health Organization

1.2. Glossary of Definitions

Abbreviations list pertains to the SAP only.

Abbreviation	Description
ACTIVE-mini	Abilities Captured Through Interactive Video Evaluation- Mini: a Microsoft-Kinect camera-based system designed to measure spontaneously or elicited movements. Through quantification of positional changes of hands and feet over time, ACTIVE-mini has the ability to measure total movement volume and speeds of movement of each extremity separately.
AE	Adverse Event: Any untoward medical occurrence in a clinical investigation subject which does not necessarily have a causal relationship with the drug or device under study.
Age	For a given event, age will be expressed in months and rounded to one decimal place. A month is standardized to a period of 30 days. Age at Event = (Date of Event – Date of Birth)/ 30. Age at Baseline = (Date of Gene Therapy Infusion – Date of Birth)/ 30.

Abbreviation	Description
Baseline	Baseline, e.g., in terms of baseline laboratory values, vital signs, or physical exam results, refers to a measurement or evaluation made prior to initiation of gene therapy infusion. If there are multiple measurements prior to the initiation of gene therapy infusion, only the latest measurement will be considered as baseline for analysis purposes.
BiPAP	Bilevel Positive Airway Pressure, a form of non-invasive mechanical pressure support ventilation.
CHOP-INTEND	Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders, a 16-item motor function assessment validated for use in infants with spinal muscular atrophy.
CMAP	Compound Muscle Action Potential, the total electrophysiological output in a muscle or muscle group recorded by surface electrodes following supramaximal stimulation of the innervating nerve.
Day 0	The day of the gene transfer.
Dose	Total vector genome delivered (in $^{vg}/_{\text{subject weight on Day } -1(\text{kg})}$).
EIM	Electrical Impedance Myography, a non-invasive technique for the evaluation of neuromuscular disease that relies upon the application and measurement of high-frequency, low-intensity electrical current.
MedDRA	Medical Dictionary for Regulatory Activities is a medical terminology used to classify adverse event information associated with the use of biopharmaceuticals and other medical products.
MUNE	Motor Unit Number Estimation, an estimate of the total number of motor units in a specific muscle or muscle group obtained through electrophysiological stimulation of muscle nerves.
NCI CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events version 4.03. The CTCAE is descriptive terminology, involving a severity scale, which is used for AE reporting in particular for unique patient populations. It is a subset of MedDRA terminology.
Permanent Ventilation	Requirement of ≥16-hour respiratory assistance per day (includes BiPAP) continuously for ≥14 days in the absence of an acute reversible illness, excluding perioperative ventilation.
SMUP	Single Motor Unit Potential, the maximal electrical output of muscle fibers of a single motor neuron.
Study Day	For any event of interest, Study Day = calendar date of event – calendar date of gene transfer. Day 0 is the Study Day of the gene transfer and any events occurring on the same calendar day as the gene transfer.
TEAE	Treatment-emergent Adverse Event = any adverse event whose onset (or worsening of an existing AE) occurred on or post day of gene therapy infusion.
Vector genome	The human SMN cDNA sequence, corresponding to the mature mRNA, cloned into the self-complementary AAV vector plasmid.

2. PURPOSE

There will be two Clinical Study Reports (CSR) written for this ongoing clinical trial:

A.) A primary CSR once all enrolled subjects have reached 15 months of age or are otherwise withdrawn from study. This CSR will be based on the primary efficacy analysis.

B.) A CSR addendum to describe safety follow-up and longer-term efficacy once all enrolled subjects have had 24 months of follow-up in the study or are otherwise withdrawn from study.

This Statistical Analysis Plan (SAP) describes the planned analysis of all data available at the time at which all patients have reached 15.0 months- of-age. This will be the point for data cut-off for all patients in the clinical study report and will include any patient data up to 24-months post treatment. The SAP provides details to guide the analyses for baseline, efficacy, and safety variables and describes the populations and variables that will be analyzed and the statistical methods that will be utilized. Analyses will be performed using SAS® Version 9.4 (SAS Institute, Inc., Cary, NC) or later under the Windows (Server 2008 R2) operating system.

This SAP supersedes the statistical considerations identified in the protocol; where considerations are substantially different, they will be so identified. This SAP is compliant with ICH E9.

2.1. Responsibilities

AveXis, Inc. is responsible for ownership and approval of the SAP.

Novella (CRO) is responsible for deriving the data set according to CDISC standards and create data set specifications based on the SAP. Novella will perform the statistical analyses and are responsible for the production and quality control of all tables, listings and figures.

2.2. Timings of Analyses

2.2.1. Quarterly Data Reviews

As of 01 April 2016, all Data Safety Monitoring Board (DSMB) activities were transferred from NCH to AveXis for overall management and obligations. The Data Safety Monitoring Board operated in advisory capacity to determine if the individual study or clinical development program for the “Phase I Gene Transfer clinical trial for spinal muscular atrophy Type 1 delivering survival motor neuron gene by self-complementary AAV9”, met one or more of the following conditions:

- clinical endpoints related to mortality or major morbidity
- evaluation of a therapy for which the safety record was not established or prior information suggesting possibility of serious toxicity with a therapy
- vulnerable patient population being studied
- *a priori* reasons for a particular safety concern.

Responsibilities of the DSMB agreed upon were:

- Review the research protocol, informed consent documents, and plans for data and safety monitoring;
- Evaluate the progress of the trial, including periodic assessments of data quality and timeliness (quarterly data cut reviews), participant recruitment, participant risk versus benefit, and other factors that can affect study outcome;
- Consider factors external to the study when relevant information becomes available, such as scientific or therapeutic developments that may have an impact on participant safety or the ethics of the trial;
- Protect the safety of the study participants;
- Review safety data to determine whether to recommend dose escalation;
- Ensure the confidentiality of the trial data and the results of monitoring.

2.2.2. DSMB Reporting and Meetings

Reports describing the status of the study were initially prepared by the Principal Investigator's staff and later AveXis (post transfer of obligations as a result of IND transfer) and sent to each DSMB member prior to each meeting for review and preparation. DSMB meetings were to occur on quarterly basis and were later aligned with the AveXis quarterly data cut reviews (as AVXS summary report completed in following month).

Reports included the following:

- A brief narrative of the study status, including the target enrollment, current and projected time to completing enrollment. Any significant events and/or difficulties should be briefly described in this narrative.
- A brief narrative for each participant describing gender, age, race and ethnicity and other relevant demographic characteristics. The narrative for each participant should briefly describe his/her study status (i.e., dose level, visit number, adverse event information).
- A timeline outlining the study progress relative to visit number for each participant, as well as time points for each SAE/Dose Limiting Toxicity. A total for Adverse Events (AEs) for each participant should be included.
- A summary of AEs by severity levels
- A listing of AE details grouped by participant
- A listing of SAE details grouped by participant
- A listing of deaths
- A summary of clinically significant laboratory test results
- Other key information related to safety/conduct of the study

2.2.3. Stopping/Discontinuation Rules

An independent Data Safety Monitoring Board (DSMB) was selected for the study. Safety data will be monitored on a continual basis throughout the trial in accordance with ICH/GCP and institutional requirements, including Sponsor Safety Management Plan (SMP). The DSMB could recommend early termination of the trial for reasons of safety. Study enrollment could be halted by the investigator and/or Sponsor when any subject experienced a Grade 3 or higher adverse event toxicity that is possibly, probably, or definitely related to the study drug. This included any subject death, important clinical laboratory finding, or any severe local complication in the injected area related to administration of the study agent.

All materials developed for the DSMB to review and all documented recommendations were maintained according to confidentiality requirements.

2.2.4. DSMB Reviews Completed

Since the initiation of this study, the DSMB has met on the following occasions:

Table 1 DSMB Meeting Dates and Decisions

Date	Decision Regarding Patient Enrollment
24Mar2014	Prior to first patient enrolled.
16Jun2014	Hold enrollment of 2nd patient until LFEs return to normal for 1st patient
10July2014	Prior to E.02's anticipated dosing on 28July2014, a follow-up meeting was held to evaluate the LFE levels of E.01. E.02 dosing was subsequently postponed due to protocol amendment for prednisolone pre-treatment.
30Oct2014	DSMB supports dose-escalation to intermediate dose.
21Apr2015	Add one more patient to 2.0 X 10^{14} vg/kg dose.
05Aug2015	Enrollment plan was for 3 younger patients at intermediate dose and 3 younger patients at high dose.
02May2016	Safety and efficacy update. Decision to terminate enrollment at Cohort 2B and not escalate to Cohort 3 had been made in October 2015.

The DSMB was provided with detailed and extensive clinical, laboratory, and immunology safety data to monitor study subject progress. During one of their meetings, with three patients having received the lowest dose, 6.7 X 10^{13} vg/kg, the decision was made to escalate from 6.7 X 10^{13} vg/kg to the next higher dose, 2.0 X 10^{14} vg/kg. Subsequent decisions were made to expand enrollment in this intermediate dose group to include a total of 12 patients, and finally in October 2015, the decision was made to terminate enrollment at 2.0 X 10^{14} vg/kg dose group and not escalate to the 3.3 X 10^{14} vg/kg dose group, bringing the final study size in both groups to n=15. As of December 2015, all 15 enrolled subjects had been treated.

On November 5, 2015, the sponsorship of the study shifted from Jerry R. Mendell, MD to AveXis, Inc. Correspondingly, a new global DSMB will be formed in the fourth quarter of 2016 and will meet quarterly to closely monitor subject safety.

3. STUDY OBJECTIVES

3.1. Primary Objective

The primary objective of this gene therapy clinical trial is the determination of safety based on the development of unacceptable toxicity, defined as the occurrence of any one Grade III or higher, unanticipated, treatment-related toxicity that requires medical treatment or presents with symptoms.

3.2. Secondary Objective(s)

The secondary objective is the determination of efficacy based upon all available data at the time when all patients have reached at least 15 month of age. Two co-primary and two co-secondary endpoints will be assessed for this purpose. One co-primary efficacy endpoint is video confirmed achievement of functional independent sitting. The other co-primary efficacy endpoint is the time from birth until death or until the patient requires ≥16 hours per day of respiratory assistance (includes non-invasive ventilatory support) continuously for ≥14 days in the absence of an acute reversible illness, excluding perioperative ventilation.

One co-secondary endpoint will include the proportion of patients who did not require parenteral nutrition prior to therapy who maintain the ability to thrive, defined by the ability to tolerate thin liquids as demonstrated through a formal swallowing test, do not receive nutrition through mechanical support (i.e., feeding tube), maintains weight (> third percentile for age and gender) at 15 months of age. The other co-secondary endpoint will include the proportion of patients who are independent of ventilatory support, defined as requiring no daily ventilator support/usage at 15 months of age, in the absence of acute reversible illness and excluding perioperative ventilation.

3.3. Additional Objectives

Additional outcomes will include the change in motor function as measured by score on the CHOP-INTEND scale from baseline and demonstration of improvement of motor function and muscle strength as determined by video confirmed achievement of significant development milestones including but not limited to the ability to maintain head control and roll over unassisted.

Bayley Scales of Infant Development© version 3 is a standardized, norm-referenced infant assessment. The gross and fine motor portions as well as speech and cognition portions of this test will administered if the child reaches or exceeds a score of 60 out of 64 on the CHOP-INTEND.

Additional exploratory outcome measures will be tested during the study as part of the program and product development plan; however, if these measurement results show efficacy and primary outcomes do not reach statistical significance, the only interpretation is that the results show a trend toward benefit but can never supersede primary measures. These exploratory outcome measures will include:

- ACTIVE-mini (Ability Captured Through Interactive Video Evaluation-mini) evaluation of infant movement ability.

- Motor neuron function will be assessed via evoked compound motor action potentials (CMAP) and motor unit number estimation (MUNE).
- Pathological status of muscles will be quantified by Electrical Impedance Myography (EIM).
- Compelling, demonstrable, documented evidence of efficacy as determined by changes in functional abilities as captured during videotaping sessions during site visits and/or captured by subject/parent/legal guardian at home.

3.4. Study Design

This is a Phase I, open-label, single injection, ascending dose, single center study to evaluate the safety and efficacy of AVXS-101 in the treatment of Spinal Muscular Atrophy Type 1.

The study will include a Screening Period of at least 30 days, a Study Period through at the time point when all patients have reached at least 15 months of age, and a post-study follow-up period up to the last patient reaches 24 months post dose (Figure 1, Study Design Schematic).

Approximately 15 clinically stable patients with AAV9 and hSMN antibody titers less than 1:50, and 2 copies of *SMN2* will be eligible to participate. The initial protocol required participants to be less than 9 months of age at time of dosing; a subsequent revision to the protocol stipulates that participants be less than 6 months of age at time of dosing.

After completion of the Day -1 visit, patients will be assigned to following 3 dosing cohorts:

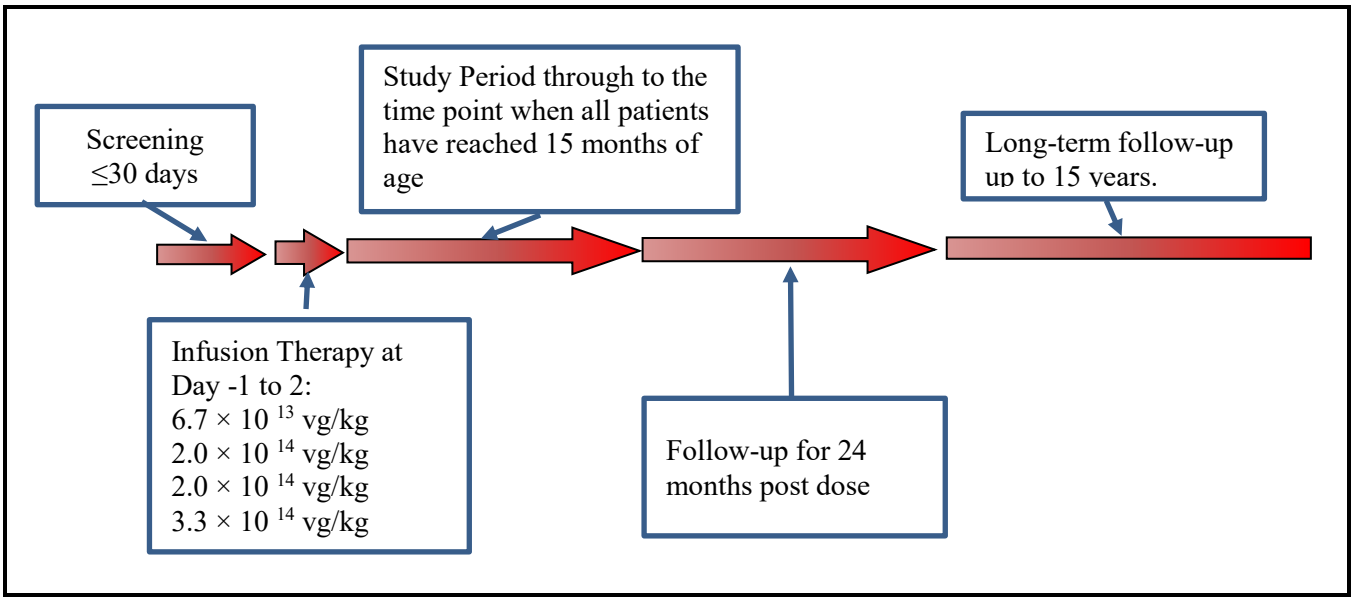
Cohort 1 (Low Dose): 6.7 X 10e13 vg/kg

Cohort 2A/2B (Intermediate Dose): 2.0 X 10e14 vg/kg

Cohort 3 (High Dose): 3.3 X 10e14 vg/kg

Cohort 1 will consist of 3 SMA patients followed by Cohort 2A with a minimum of 3 and no more than 6 SMA patients to receive a single infusion at escalating doses. Cohort 2B and Cohort 3 will each consist of 3 SMA patients to receive a single infusion at escalating doses. Cohort 2B will be done first and if efficacy is favorable but does not fully improve subjects to normal, Cohort 3 at high dose will be implemented.

Figure 1 Study Design Schematic



By mutual decision of the investigator, Sponsor and DSMB in October 2015, enrollment was terminated with the last subject in Cohort 2A/2B and Cohort 3 was never initiated.

Table 2 Schedule of Assessments

Study Interval	Baseline	Vector Infusion				Follow Up Year 1									Follow Up Year 2	
	Screening	(Inpatient)				(Outpatient)										
Visit	1	2				3	4	5	6	7	8	Monthly Visit 9 (Month 4), 10 (Month 5), 12 (Month 7), 13 (Month 8), 15 (Month 10), 16 (Month 11)	Every 3 Months Visit 8 (Month 3), 14 (Month 9)	Every 6 Months Visit 11 (Month 6) , 17 (Month 12)	Every 3 Months Visit 20 (Month 15), 26 (Month 21)[g]	Every 6 Months Visit 23 (Month 18), 29 (Month 24)
Days in Study[a]	-30	-1	0	1	2	7	14	21	30	60	90	Up to 12 Months			13-24 Months	
	(± 7)														+/- 21 days	
Informed Consent	X															
Chest X-Ray	X															
Medical History	X															
Physical Exam + Vitals	X	X[c]	X[d]	X	X	X	X	X	X	X	X	X	X	X	X	X
Pulmonary Assessment	X			X		X	X	X	X	X	X	X	X	X	X	X
Photograph Injection site	X	X	X	X	X	X	X	X	X	X	X	X	X	X		
Swallowing Test	X													X		X
Pulse Oximetry[d]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Capillary Blood Gas	X	X	X	X	X											

Study Interval	Baseline	Vector Infusion				Follow Up Year 1									Follow Up Year 2	
	Screening	(Inpatient)				(Outpatient)										
Visit	1	2				3	4	5	6	7	8	Monthly Visit 9 (Month 4), 10 (Month 5), 12 (Month 7), 13 (Month 8), 15 (Month 10), 16 (Month 11)	Every 3 Months Visit 8 (Month 3), 14 (Month 9)	Every 6 Months Visit 11 (Month 6) , 17 (Month 12)	Every 3 Months Visit 20 (Month 15), 26 (Month 21)[g]	Every 6 Months Visit 23 (Month 18), 29 (Month 24)
Days in Study[a]	-30	-1	0	1	2	7	14	21	30	60	90	Up to 12 Months			13-24 Months	
	(± 7)														+/- 21 days	
Safety Labs (Blood)	X	X		X	X	X	X	X	X	X	X		X	X	X	X
Coagulation Studies (PT/INR/PTT)											X		X	X	X	X
Safety Labs (Urinalysis)	X	X		X	X	X	X		X		X		X	X	X	X
Immunology (AntiAAV9/SMN Ab & T-Cells)	X					X	X	X	X	X	X		X	X	X	X
Research Blood[e]	X							X	X	X	X		X	X	X	X
CHOP-INTEND[f] (with video)	X	X							X	X	X	X	X	X	X	X
Bayley[b]	X	X							X	X	X	X[b]	X[b]	X[b]	X	X
ACTIVE-mini	X	X				X	X	X	X	X	X	X	X	X	X	X
CMAP/MUNE/EIM	X								X		X		X	X	X	X
Research Urine	X			X	X	X	X	X	X	X	X	X	X	X	X	X
Research Saliva & Stool	X			X	X	X	X	X	X	X	X	X	X	X	X	X
Development Milestones/Gross Motor Skills Checklist (with video)	X								X					X	X	X
ECHO/ECG	X								X					X		X
Prednisolone Dosing		X	X	X	X	X	X	X	X							
Gene Transfer			X													
Adverse Events	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Concomitant Medications	To be collected from 2 weeks before study dose until final study visit, recorded on separate CRF															

a Visits on Days 7, 14 and 21 allow a window of ±2 days; all monthly visits following allow a window of ±7 days

b The gross and fine motor subtests will be administered monthly through the timepoint that the patient reaches 13.6 months of age or 12 months post-dose, whichever is later, if a patient reaches or exceeds a score of 60 out of 64 on the CHOP-INTEND. The language and cognition subtests will be administered every three months if a patient reaches or exceeds a score of 60 out of 64 on the CHOP-INTEND. CHOP-INTEND assessment will be discontinued and only the Bayley will be administered for subjects that achieve two consecutive scores of ≥62.

c Vital signs recorded every four (4) hours during inpatient hospitalization.

d Continuous monitoring during gene transfer procedure. Axillary temperature to be captured pre- and post- infusion.

e Research blood sample will be used to perform baseline exon 7 modification testing and could also be used to re-confirm SMA Type 1 diagnosis, SMN2 copy number, and exon-7 modification testing through a third-party laboratory.

f Subjects that achieve two consecutive scores of ≥62 may cease further CHOP-INTEND assessments, as per PI, physical therapist, and sponsor decision.

g Quarterly visits in Year 2 will have a +/- 21 day window. Missed visits must be made up within the allowed window; there must never be 6 months between study visits.

3.5. Subject Selection

The current version of the protocol (v14.0) allows for subjects 6 months of age or younger at time of gene therapy infusion diagnosed with SMA Type 1 and possessing 2 copies of *SMN2* without c.859G>c modification in exon 7. Previous versions of the protocol allowed for subjects as old as 9 months of age at time of infusion. See protocol for details.

3.5.1. Inclusion Criteria

See protocol for details.

3.5.2. Exclusion Criteria

See protocol for details.

3.6. Determination of Sample Size

This is a Phase I dose-escalation trial, with safety as the primary objective. Sample size was not determined through statistical justification. Based on mutual decision of investigator, Sponsor and DSMB, total 15 eligible patients were enrolled to 2 cohorts.

3.7. Treatment Assignment and Blinding

This is an open-label study. All enrolled subjects received AVXS-101.

3.8. Administration of Study Medication

Refer to Section 3.4 of the SAP.

3.9. Study Procedures and Flowchart

Refer to Section 3.4 of the SAP.

3.10. Statistical Hypotheses

3.10.1. Primary Efficacy Hypothesis

For the co-primary efficacy endpoint of functional independent sitting and the other co-primary efficacy endpoint of survival rate, a hierarchical testing procedure will be used to control Type I error rate across the primary objectives ($\alpha = 0.025$, one-sided). The superiority hypothesis of functional independent sitting compared to natural historical rate 0.1% will be tested first at a significance level of 2.5%.

$$\mathbf{H_0}: p_{AVXS-101} = 0.1\%$$
$$versus\ the\ alternative$$
$$\mathbf{H_a}: p_{AVXS-101} > 0.1\%,$$

where p is the proportion achieving functional independent sitting.

The co-primary of survival will be assessed by superiority hypothesis testing of the survival rate compared to the 25% survival rate observed from historical data (Finkel et al, 2014) at a significant level of 2.5%.

$\mathbf{H_0}$: $p_{AVXS-101} = p_{HISTORICAL-FINKEL}$
versus the alternative
$\mathbf{H_a}$: $p_{AVXS-101} > p_{HISTORICAL-FINKEL}$,
where p is the proportion surviving event-free to 13.6 months of age.

However, the comparison of survival rate between AVXS-101 and historic data (Finkel et al, 2014) will only be confirmed as statistically significant if this test is significant on both mITT set and the full analysis set and the preceding test for superiority of functional independent sitting is statistically significant.

The comparisons will be summarized by a point estimate of the difference with a 97.5% confidence interval.

3.10.2. Secondary and Additional Efficacy Hypotheses

One-sided superiority test will be performed to test the following hypotheses at a significance level of 2.5% for co-secondary endpoints and at level of 5% for additional efficacy endpoints.

$\mathbf{H_0}$: $p_{AVXS-101} = 0.1\%$
versus the alternative
$\mathbf{H_a}$: $p_{AVXS-101} > 0.1\%$,
where p is the proportion of maintaining ability to thrive / independent of ventilatory support / attaining the milestone / cut-off threshold of CHOP-INTEND score during the observation period.

The comparisons will be summarized by a point estimate of the difference with a 97.5% confidence interval for co-secondary endpoints and 95% confidence interval for additional efficacy endpoints.

4. ENDPOINTS

4.1. Safety Endpoint

The primary safety endpoint is the development of unacceptable toxicity, defined as the occurrence of any one Grade III or higher, unanticipated, treatment-related toxicity that presents with clinical symptoms and requires medical treatment.

4.2. Efficacy Endpoints: (see Section 8 for Details)

4.2.1. Primary Efficacy Endpoints

One co-primary efficacy endpoint is achievement of functional independent sitting. The other co-primary efficacy endpoint is survival which is defined as time from birth to either (a) death or (b) requirement of ≥16-hour respiratory assistance per day (includes non-invasive ventilation) continuously for ≥14 days in the absence of an acute reversible illness, excluding perioperative ventilation.

4.2.2. Secondary Efficacy Endpoints

One co-secondary outcome will include the proportion of patients who did not require parenteral nutrition prior to therapy who maintain the ability to thrive, defined by the ability to tolerate thin liquids as demonstrated through a formal swallowing test, do not receive nutrition through mechanical support (i.e., feeding tube), maintains weight (> third percentile for age and gender) at 15 months of age. The other co-secondary secondary outcome will include the proportion of patients who are independent of ventilatory support, defined as requiring no daily ventilator support/usage at 15 months of age, in the absence of acute reversible illness and excluding perioperative ventilation.

4.3. Additional Efficacy Endpoints: (see Section 8 for Details)

- Development of Significant Motor Function Milestones.
- Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders (CHOP-INTEND).
- Bayley Scales of Infant Development© version 3. This test will be administered if the child reaches or exceeds a score of 60 out of 64 on the CHOP-INTEND.
- ACTIVE-mini (Ability Captured Through Interactive Video Evaluation-mini) evaluation of infant movement ability.
- Evoked compound motor action potentials (CMAP) and motor unit number estimation (MUNE).
- Electrical Impedance Myography (EIM).

4.4. Pharmacokinetic Endpoints

Not applicable for this SAP.

4.5. Pharmacodynamic Endpoints

Not applicable for this SAP.

4.6. Additional Safety Endpoints

Additional safety analyses include:

4.6.1. Adverse Events

Other than Grade III or higher adverse events used to define unacceptable toxicity in the primary endpoint, all adverse events will be assessed for their seriousness, relatedness to study treatment, relationship to study discontinuation and severity according to CTCAE version 4.03 criteria.

4.6.2. Vital Signs

Blood pressure, respiratory rate, pulse, and axillary temperature will be collected at every visit. For detailed schedule of events table, see Table 2.

4.6.3. Physical Examination

Physical examination will include review of the following systems: Head, eyes, ears, nose, throat (HEENT), lungs/thorax, cardiovascular, abdomen, musculoskeletal, neurologic, and genitourinary.

4.6.4. Laboratory Evaluations

Serum will be collected throughout the study for standard blood chemistry, hematology, and coagulation tests. Urine will be collected throughout the study for standard urinalysis exams.

4.6.5. Usage of Non-oral Feeding Support

A swallowing test will be performed at baseline and every 6 months to determine if the subject has signs of aspiration. If the test is positive for aspiration, the patient will be instructed to use an alternative method to oral feeding. Once implanted, a non-oral method of feeding support may later be removed. For each placement or removal event, the type of support (type of tube), date of placement, and date of removal will be noted. Actual use of non-oral feeding support will be quantified through the recording of volume, frequency of use, duration, and calories. Actual usage information was not requested or captured in source data from the start of the study, so these data will not be complete for all study subjects.

4.6.6. Pulse Oximetry

Pulse oximetry will be measured throughout the study through a small infrared light attached to the end of the patient's finger.

4.6.7. Cardiovascular Safety Evaluations

Echocardiograms and electrocardiograms will be conducted at screening and every 6 months.

4.6.8. Immunology

Immunoreactivity to AAV9 and SMN will be measured in antibody titer levels (in 2-fold serial dilutions) as determined by enzyme-linked immunosorbent assays (ELISAs). T-cell response to AAV9 and SMN will be measured in number of spot forming cells per million peripheral blood mononuclear cells (PBMCs) as determined by Enzyme-Linked ImmunoSpot (ELISpot) assays.

4.6.9. Concomitant Medications

Prior and concomitant medications will be captured in the eCRF from two weeks prior to study dosing through the last study visit and coded using the WHODrug dictionary, version March 2016 C Final 2.

4.7. Health-economics Endpoints

Not applicable for this SAP.

4.8. Other Endpoints

Additional analyses of the two co-primary and other additional endpoints utilizing a matched case control/natural history cohort as a comparison may be incorporated, as appropriate.

5. ANALYSIS SETS

Safety will be assessed using the combined experience of Cohort 1 and Cohort 2. Efficacy results will be presented separately for Cohort 1 and Cohort 2, with emphasis on results in Cohort 2, as this is the intended dose. In addition, efficacy results will also be presented in efficacy subgroups and mITT set as well.

5.1. Safety Set

All subjects who underwent gene therapy infusion will be included.

5.2. Intent-to-Treat or Full Analysis Set

5.2.1. Intent-to-Treat (ITT)

Any subject who underwent gene therapy infusion will be included.

5.2.2. Full Analysis Set (FAS)

Any subject who underwent gene therapy infusion and had at least one post-infusion visit will be included.

5.3. Modified Intent-to-Treat

The modified ITT (mITT) population is a subset of the ITT population, which will consist of all patients with bi-allelic deletions of SMN1 and baseline CHOP-INTEND scores ≥20 who received an infusion treatment of AVXS-101 at dosing 2.0×10^{14} vg/kg. Efficacy analyses, except for the co-secondary endpoint of "ability to thrive" (see Section 5.7) will be conducted on the mITT population.

5.4. Per Protocol or Evaluable Set

The data from the ITT population in Cohort 2 will also be analyzed according to dosing age as per protocol (Amendment Addendum 1).

5.4.1. Efficacy Evaluable Set (EES)

Any subject who remains in the study beyond 13.6 months of age or whose death or permanent ventilation status can be ascertained at age 13.6 months of age will be included. Early withdrawals from the study for reasons other than death or permanent ventilation will not be included in the EES.

5.5. Pharmacokinetic Set

Not applicable for this SAP.

5.6. Pharmacodynamic Set

Not applicable for this SAP.

5.7. Other Analysis Sets

The Ability to Thrive ITT population is a subset of the ITT population, which will consist of all patients with bi-allelic deletions of SMN1 and baseline CHOP-INTEND scores ≥20 who received an infusion treatment of AVXS-101 at dosing 2.0×10^{14} vg/kg and who did not require parenteral nutrition prior to administration of AVXS-101. The efficacy analysis for the co-secondary outcome of "ability to thrive" will be conducted on the Ability to Thrive ITT population.

5.8. Protocol Deviations

As this is a single dose study, protocol deviations will be described; however, there will be no separate statistical analysis.

6. GENERAL ASPECTS FOR STATISTICAL ANALYSIS

6.1. General Methods

All hypothesis testing will be conducted at the 0.05 level of significance (α=0.05), but primary and secondary efficacy analyses, which will be conducted at the 0.025 level of significance. Tests may be one-sided or two-sided, as appropriate. Categorical measures, such as percent surviving event-free, will be summarized using count and percentages.

Continuous data, such as CHOP-INTEND scores, will be summarized using count, mean, median, standard deviation (SD), minimum, and maximum. For continuous data specified to be analyzed using parametric procedures, non-parametric procedures will be used if the parametric procedure is felt to be inappropriate.

Summary tables, listings, and figures, and statistical analyses will be done using SAS version 9.3 or higher or R, version 2.0 or higher.

Efficacy analyses will be conducted on the Full Analysis Set or the Efficacy Evaluable Set or mITT set or Ability to Thrive ITT set, as indicated. Safety analyses will be conducted on the Safety Analysis Set only.

When an analysis calls for statistical modeling of the outcome variable using potential confounders as well as the treatment effect, the analysis will begin with the full model. If a potential confounder is not statistically significant in the full model, it will be dropped and the analysis repeated using a reduced model.

As noted earlier, the primary analyses for efficacy will be assessed when all patients have reached at least 15 months of age (15 months-of-age data cut-off). A follow-up safety analysis will be completed at the time point at which the last patient reaches 24 months post-dose. Since visits are scheduled relative to time in the study, at the time of the 15 months-of-age data cut-off, the earliest enrolled subjects will have been in the study much longer and will have more follow-up visits than the most-recently enrolled subjects. Some of the long-term follow-up visits (e.g., 21 months and 24 months post-infusion) will have too few subjects at the time of the 15 months-of-age data cut-off to warrant tabular summaries. For these extreme time points, data may be presented in listings only.

Despite this limitation, for the remainder of this section, analyses will be described according to the Schedule of Assessments detailed in the protocol.

6.2. Key Definitions

6.2.1. Definition of Baseline

The baseline value refers to the last non-missing measurement collected before gene therapy infusion of study drug. On Study Day 0, all assessments should be performed prior to administering the gene therapy infusion of study drug per protocol. The baseline value is therefore determined by the last non-missing measurement collected on or before the first day of study drug administration. If multiple measurements are recorded on the same day, the last measurement recorded prior to the gene therapy infusion will be used as the baseline value. If these multiple measurements occur at the same time or time is not available, then the average of

these measurements (for continuous data) or the worst among these measurements (for categorical data) will be considered as the baseline value. This same baseline value will be used for the treatment and post-treatment periods.

6.2.2. Definition of Study Days (Days Relative to the Gene Therapy Infusion)

Study Days are calculated for each time point in the treatment period relative to the gene therapy infusion of study drug. Study Days are negative values when the time point of interest is prior to the study drug infusion day. Study Days are positive values when the time point of interest is after the first study drug dose day. Study Day 0 is the day of the infusion of the gene therapy study drug.

6.2.3. Definition of Final Treatment Value

The final treatment value for each subject is the last non-missing measurement collected after Study Day 1 and on or before the date when the all patient have had a visit at or after 15 months of age.

There is no upper bound for the "15 months" age, that is, no visit is too late to represent the "15 months of age" analysis. The earliest visit on or after 450 days of age will be used to represent the "15 months of age" analysis.

For the purposes of the co-primary endpoint of survival (see 8.1.2), the first non-missing value after the participant reaches 13.6 months of age will serve as the final treatment value. For the co-secondary endpoints of "maintain ability to thrive" (see 8.2.1) and "independence of ventilatory support" (see 8.2.2) the 15 month of age visit, as defined above, will serve as the final treatment value. For all other endpoints, including the co-primary endpoint of functional independent sitting, the analysis will include all measurements for a given patient collected on or before the date when all patients in the study have had a visit at or after 15 months of age. For most patients, data will include measurements obtained at ages older than 15 months of age.

6.2.4. Definition of Final Post-Treatment Value

The final post-treatment value for each subject is the last non-missing measurement collected.

6.2.5. Derived and Transformed Data

6.2.5.1. Primary Safety and Efficacy Variables

The Primary safety endpoint is defined as if the patient meets following criteria:

- Unanticipated toxicity event of Grade ≥3 (as per CTCAE), *and*
- 'Possibly Related' or 'Probably Related' or 'Definitely Related' to study treatment, *and*
- Requires medical treatment or presents with symptoms

This will be collected at all scheduled visits (see Table 3) from subject's CRF.

The co-primary efficacy endpoint is functional independent sitting which is defined as:

- Functional independent sitting: this is assessed by video evaluation by an expert reviewer of videos taken either at scheduled visits (see Table 4) or provided by parent/legal guardian, if patient meets the criterion: maintain a sitting position independently for at least 30 seconds. This will be collected from external dataset.

The other co-primary endpoint is defined as the time from birth date to any one of the following events whichever occurred first:

- Death, *or*
- Invasive ventilatory support, *or*
- Noninvasive Ventilatory support ≥16-hour per day *and* continuously ≥14 days

The primary endpoint variables will be collected at all scheduled visits (see Table 3) from subject's CRF.

6.2.5.2. Supportive (Secondary and Additional) Efficacy Variables

Supportive (secondary) efficacy variables are as follows:

- Proportion of patients who did not require parenteral nutrition prior to therapy who maintain the ability to thrive, defined by the ability to tolerate thin liquids as demonstrated through a formal swallowing test, do not receive nutrition through mechanical support (i.e., feeding tube), maintains weight (> third percentile for age and gender) at 15 months of age.
- Proportion of patients who are independent of ventilatory support, defined as requiring no daily ventilator support/usage at 15 months of age, in the absence of acute reversible illness and excluding perioperative ventilation.

Supportive (additional) efficacy variables are as follows:

- Motor Function Milestones, Gross Motor Skill: these are also assessed by video evaluation and collected from external dataset at scheduled visits (see Table 4). These milestones will include two additional alternative (secondary) definitions for the milestone of sitting without support: 1) as defined by the WHO (Child sits up straight with head erect for at least 10 seconds; child does not use hands or arms to balance body or support position); and 2) as defined by Nationwide Children's Hospital – sitting without support for 15 or more seconds.
- CHOP-INTEND summary score: This is assessed by evaluating the Baseline CHOP-INTEND scores compared to each scheduled visit time points (see Table 4), collecting from subject's CRF. A summary score is obtained by summing the individual best-side scores.
- Achievement of Selected Threshold Scores on CHOP-INTEND: clinically relevant threshold scores were setup for categorizing CHOP-INTEND summary scored: ≥40, ≥50, ≥60.
- Bayley Scales of Infant Development©: this is administered to patients whose CHOP-INTEND scores of ≥60. Five domains of cognition, language, social-emotional, motor and

adaptive behavior will be assessed from subject's CRF at scheduled visits (see Table 4) with both raw and scaled scores provided.

- ACTIVE-mini: this is video evaluation to measure spontaneous or elicited movements. But it is still under development, only for exploratory examination.
- CMAP / MUNE: CMAP is an electromyographical measure of muscle function and is collected from subject's CRF. MUNE calculation see Section 8.2.9 and Appendix 21.4.

6.2.6. Completers

Any subject who has a baseline and completes the final study visit (reaches 15 months of age following gene transfer) shall be considered as having completed the study. Any subject who has a baseline and completes the final follow-up visit (24 months following gene transfer) shall be considered as having completed the post-treatment follow-up.

6.3. Missing Data

6.3.1. CHOP-INTEND and Partial Assessments

The CHOP-INTEND is an instrument composed of 16 body-movement and body-control tasks to quantify motor function in infants, with a minimum score of 0 and a maximum score of 64. The instrument has been studied and validated in infants with spinal muscular atrophy. However, there are circumstances, as discussed below, when a complete assessment is not possible. The scoring of the CHOP-INTEND proceeds differently under those circumstances:

- Major, externally-imposed, systemic limitation of movement.
- When there is a major, systemic, and externally imposed limitation that would prevent the accurate assessment of multiple CHOP-INTEND items, the total CHOP-INTEND score is regarded as missing. A partial assessment may be conducted, but one or more items will be scored as "Could Not Test". An example has been an infant who was in a body cast due to scoliosis. A total CHOP-INTEND score will not be computed when one or more of the items are scored as "Could Not Test".
- Individual CHOP-INTEND items in which the physical therapist and the parent/guardian agree that attempting to perform the task would be unsafe and that the subject will be incapable of scoring any points in the evaluation.
- In that case, the individual task is not performed and the item score is imputed as a 0. The total CHOP-INTEND score is obtained by summing all of the items, including the items whose score was imputed as 0. An example has been an infant who was too big to assess the Galant response safely.

6.3.2. Death or Permanent Ventilation Status in Intent-to-Treat Efficacy Analysis

One efficacy endpoint is the Time from Birth to Death or Permanent Ventilation. As described below, this endpoint will be analyzed as the proportion of subjects surviving without requiring permanent ventilation to 13.6 months of age.

For analysis of this endpoint using the ITT population, subjects who terminate from the study prior to reaching 13.6 months of age without requiring permanent ventilation before study termination will be regarded as treatment failures.

For analysis of this endpoint using the EES population, only subjects whose death/permanent ventilation status at 13.6 months of age is known will be used.

6.3.3. All Other Data

No imputation for missing data will take place for any other data collected.

6.4. Visit Windows

For efficacy and safety analysis, the time windows specified in Table 3 to Table 7 describe how efficacy and safety data assigned to protocol-specified time points during treatment and post-treatment periods, respectively. All time points and corresponding time windows are defined based on the gene therapy infusion date.

If more than one assessment is included in a time window, the assessment closer to the nominal time will be used. If there are two observations equally distant to the nominal time, the latest one will be used in analyses.

If multiple measurements are made on the same day for a safety laboratory parameter or a vital sign parameter, the average of the values will be used in analysis. For summaries of shifts from baseline and potentially significant values, multiple values on the same day will not be averaged; all values will be considered for these analyses.

Table 3 Analysis Time Windows for Dose Limiting Toxicity, Pulse Oximetry, Vital Signs, and Laboratory

Scheduled Visit	Nominal Days (Study Day)	Acceptable Analysis Window in Study Days (Min Day – Max Day)
Screening	-30	-37 to -23
Baseline	-	<0
Day 0	0	0
Day 1	1	1
Day 2	2	2
Week 1	7	2-11
Week 2	14	12-17
Week 3	21	18-24
Month 1	30	25-45
Month 2	60	46-75
Month 3	90	76-105
Month 4	120	106-135
…	…	…
Month K	K*30	K*30-14, K*30+15

Table 4 Analysis Time Windows for CHOP-INTEND, Bayley and Video Assessments

Scheduled Visit	Nominal Days (Study Day)	Acceptable Analysis Window in Study Days (Min Day – Max Day)
Baseline	-	≤0
Month 1	30	1-45
Month 2	60	46-75
Month 3	90	76-105
Month 4	120	106-135
…	…	…
Month K	K*30	K*30-14, K*30+15
Final treatment visit		1 to ≤2 days after last patient reaches 15 months of age (450 days)

Table 5 Analysis Time Windows for ACTIVE-mini

Scheduled Visit	Nominal Days (Study Day)	Acceptable Analysis Window in Study Days (Min Day – Max Day)
Baseline	-	≤0
Week 1	7	1-11
Week 2	14	12-17
Week 3	21	18-24
Month 1	30	25-45
Month 2	60	46-75
Month 3	90	76-105
Month 4	120	106-135
…	…	…
Month K	K*30	K*30-14, K*30+15

Table 6 Analysis Time Windows for CMAP/MUNE/EM

Scheduled Visit	Nominal Days (Study Day)	Acceptable Analysis Window in Study Days (Min Day – Max Day)
Baseline	-	≤0
Month 1	30	1-59
Month 3	90	60-134
Month 6	180	135-224
…	…	…
Month K	K*30	K*30-45, K*30+44

Table 7 Analysis Time Windows for Urinalysis

Scheduled Visit	Nominal Days (Study Day)	Acceptable Analysis Window in Study Days (Min Day – Max Day)
Screening	-30	-37 to -23
Baseline	-	≤0
Day 1	1	1
Day 2	2	2
Week 1	7	2-11
Week 2	14	12-17
Month 1	30	25-45
Month 3	90	76-105
Month 6	180	166-195
…	…	…
Month K	K*30	K*30-14, K*30+15

6.5. Pooling of Centers

This is a single-center study, so pooling is not applicable.

6.6. Subgroups

2 Safety Subgroups will be assessed as part of the analysis:

- No Prednisolone Pre-Treatment – This subgroup will include subject E-01, the first study participant, who did not receive pre-treatment with steroids prior to administration of AVXS-101.
- Prednisolone Pre-Treatment – This subgroup will include subjects who receive pre-treatment and post-treatment maintenance with oral steroids to manage elevated LFE and T-cell response to AVXS-101.

3 Efficacy Subgroups will be assessed as part of the analysis:

- Efficacy subgroup 1: Subjects treated as part of the first dosing cohort ($6.7 * 10^{13}$ vg/kg)
- Efficacy subgroup 2: Subjects treated as part of the second dosing cohort ($2.0 * 10^{14}$ vg/kg) who were less than 6 months of age at the time of dosing
- Efficacy subgroup 3: Subjects treated as part of the second dosing cohort ($2.0 * 10^{14}$ vg/kg) who were between 6 and 9 months of age at the time of dosing

7. DEMOGRAPHIC, OTHER BASELINE CHARACTERISTICS AND MEDICATION

7.1. Subject Disposition and Withdrawals

The number and percent of subjects who are still in the study when all patients reach 15 months-of-age will be tabulated by cohort.

Among those who discontinued the study prior to when the patient reaches his or her 15 months-of-age visit, the distribution of reasons for discontinuation will be enumerated by cohort.

The number and percentage of screened subjects who screen failed and the reasons for screen failure (inclusion/exclusion criteria, withdrew consent, and/or other) will be summarized in a table. A CSR listing of reason for screen failure will be provided for all subjects who screen failed.

Reasons for dose modifications will be presented in the CSR listings.

This analysis will be conducted on the Safety Analysis Set.

7.2. Demographic Data

The age of the subject at the time of gene therapy infusion will be summarized by cohort. The distribution of subjects by sex, ethnicity (Hispanic/Latino vs. Non-Hispanic/non-Latino), and race will be presented. Patient demographics will be summarized using the Safety Analysis Set.

Demographic data will be determined using the following calculations:

Age at Study day 1 = (Study day 1 visit date - date of birth + 1) / 365.25, expressed in months.

Height (in cm) = height (in inches) * 2.54

Weight (in kg) = weight (in lbs) * 0.4536

The following statistics regarding the subject's characteristics at birth will be summarized: gestational age (weeks), birth weight (kg), birth length (cm), and head circumference (cm).

The presence of significant medical conditions obtained from medical history will summarized by cohort. In particular, the following parameters will be summarized regarding symptoms and history of Spinal Muscular Atrophy: age at symptom onset, baseline SMA symptoms, family history of SMA, and number of siblings affected by SMA.

Not all of these baseline characteristics were collected at the start of the study, so some subjects may have incomplete data. Patient baseline characteristics will be summarized on the Safety Analysis Set.

7.3. Medical History and Concomitant Diseases

Medical history data will be summarized and presented using body systems and conditions/diagnoses as captured on the eCRF. The body systems will be presented in alphabetical order and the conditions/diagnoses will be presented in alphabetical order within each body system. The number and percentage of subjects with a particular condition/diagnosis

will be summarized for each treatment group and overall. Subjects reporting more than one condition/diagnosis within a body system will be counted only once for that body system.

7.4. Prior and Concomitant Medications

Prior and concomitant medications will be summarized by cohort and overall. A prior medication is defined as any medication taken prior to the date of the gene therapy infusion of study drug. A concomitant medication is defined as any medication that started prior to the date of the infusion of study drug and continued to be taken after the infusion of study drug or any medication that started on or after the date of the infusion of study drug. The number and percentage of subjects taking prior or concomitant medications will be summarized by generic drug name based on the WHO Drug Dictionary.

7.4.1. Specific Medication Subgroups

In an attempt to reduce the host immune response to the AAV-based therapy, patients received prophylactic prednisolone (approximately 1 mg/kg/day) 24 hours prior to the gene transfer and continuing for approximately 30 days on that dose with the following guidelines for tapering thereafter: When AST and ALT exceed 120 IU/L, prednisolone will be maintained until enzymes fall below this level while at the same time monitoring T-cell response for decreases below 50 SFC per 10^6 PBMCs to Day 30 (see protocol for details).

This management was agreed upon with FDA in the revised protocol submitted in an IND update 31 Oct 2014, Serial No. 004.

The total number of days receiving prednisolone and total cumulative dose of prednisolone administered during the entire study (mg/kg) will be computed for each subject.

To compute total cumulative dose, the total dosing period is subdivided into dosing intervals represented by constant dose levels. On the day of a dosage change, the entire day is represented under the new dosing interval at the new dose.

For example, consider a subject who receives 1.0 mg/kg of prednisolone for Day 1 to Day 20, then on Day 20, dose is elevated to 2.0 mg/kg until Day 30, at which point it is tapered to 1.5 mg/kg. On Day 35, the dose is lowered to 1.0 mg/kg and continues until Day 40 when it is lowered further to 0.5 mg/kg until all prednisolone dosing stops on Day 45. For this subject,

Total Cumulative Dose= (1.0 mg/kg x 19 days) +
(2.0 mg/kg x 10 days) +
(1.5 mg/kg x 5 days) +
(1.0 mg/kg x 5 days) +
(0.5 mg/kg x 6 days) = 54.5 mg/kg

Exposure will be summarized for the Safety Analysis Set.

7.4.2. Other Therapies

Not applicable for this SAP.

8. EFFICACY

8.1. Primary Efficacy Endpoints and Analysis

8.1.1. Co-Primary Efficacy Endpoint: Functional Independent Sitting

An independent expert reviewer will evaluate videotapes of physical examinations, CHOP-INTEND or Bayley assessments conducted at clinic visits. Parent(s)/legal guardian(s) may also share home videos demonstrating achievement of functional abilities and the reviewer may evaluate those videos for evidence of milestone achievement. The reviewer will judge whether the video reveals evidence of the milestone achievement of functional independent sitting (can maintain a sitting position independently for at least 30 seconds).

The number and percent of subjects whom, through video evidence, exhibit the milestone achievement by the time of the Efficacy Data Cut-off will be summarized by cohort. The observed proportion attaining the milestone during the Observation Period will be compared to zero using a one-sided Exact Binomial Test. To make computation of the *p*-value possible, the value of 0.1% will be used in place of a literal zero.

These analyses will be performed on the Full Analysis Set and mITT set.

8.1.2. Co-Primary Efficacy Endpoint: Time from Birth to Death or Surrogate for Death (Permanent Ventilation)

The co-primary efficacy endpoint is time from birth to either (a) death or (b) permanent ventilation, defined as requirement of invasive ventilation or ≥16–hour respiratory assistance per day (includes non-invasive ventilatory support) continuously for ≥14 days in the absence of an acute reversible illness, excluding perioperative ventilation. Permanent ventilation, so defined, is considered a surrogate for death.

An “acute reversible illness” is defined as any condition other than spinal muscular atrophy that results in increased medical intervention (i.e., increased requirement for external ventilator support; use of other concomitant meds as rescue) requirements and is expected to be reversible or improved following definitive intervention (i.e., surgery, antibiotics) or introduction of escalated supportive care, such as hospitalization (i.e., for upper respiratory infection, spontaneous fracture). The specific duration of the condition antecedent intervention shall not be considered in the definition of “acute”. The date of “definitive intervention” shall be defined as the date of provision of a procedure (i.e., surgery, etc.) or medication (i.e., antibiotics) intended to cure or substantially improve the condition. For conditions such as viral respiratory infections for which supportive care is provided, the date of “definitive intervention” shall be considered the date of hospitalization or substantial escalation of care.

“Perioperative” use reflects any alteration of ventilator use related to a surgical or other medical procedure of any nature for which the participant received medications that could impair or interfere with respiratory function.

For a participant who develops an acute reversible illness and/or requires perioperative ventilator support, a recovery period not to exceed 21 days following the date of definitive intervention will be instituted. Following this recovery period, the condition will be considered sub-acute and the

participant will become evaluable with regards to the surrogate survival endpoint (Requirement of ventilator support of ≥16 hours/day for greater than 14 days).

Example: Using this approach it would mean that on day 0, subject A receives definitive intervention for an acute reversible illness resulting in ventilator support for ≥16 hours/day. Days 1-21 will be provided to permit recovery from the acute reversible illness. On Day 22, the participant is no longer considered to have an acute illness. Should the participant continue to require ≥16 hours/day of ventilator support from day 22 to day 36, he or she shall be considered to meet the surrogate endpoint.

Statistical approach: Ordinarily, such an endpoint would suggest estimating the permanent-ventilation-free survival function. However, as of the latest data available at the time of this SAP (31May2016), out of the 15 subjects treated, 9 were enrolled long enough ago to have been observed to 13.6 months of age. All 9/9 (100%) reached 13.6 months of age without requiring permanent ventilation. Anticipating this continued level of success until the efficacy data cut-off, a typical survival analysis would be impractical. Therefore, the primary analysis of the co-primary endpoint is re-expressed as a proportion: Did the subject survive without requiring permanent ventilation to 13.6 months of age: (Yes/No)?

The proportion surviving event-free to 13.6 months of age will be computed for Cohort 1 and Cohort 2 in the Intent-to-Treat analysis population. Subjects who terminate the study prior to reaching 13.6 months of age for any reason will be considered treatment failures.

As a comparator, in a natural history study of SMA Type 1 subjects, Finkel (2014) estimated that only 25% of SMA Type 1 subjects with 2 copies of *SMN2* would survive event-free to 13.6 months of age.

The observed proportion surviving in the current study will be compared to the natural history estimate of 25% using a one-sample Exact Binomial test. This comparison will be done separately for Cohort 1 and Cohort 2.

If status at 13.6 months of age is coded using variable SURVIVE as 1 if subject survived to 13.6 months of age without requiring permanent ventilation, and 2 otherwise, then the SAS code to generate the endpoint analysis is:

```
proc format;
        value survf
        1= 'Yes'
        2= 'No'
        ;
run;
proc sort data=x ;
        by cohort survive ;
run;
proc freq data=x  order=data;
        by cohort;
```

```
exact binomial;
format survive survf. ;
tables survive /binomial( wilson p=.25) ;
run;
```

These analyses will be repeated on the Efficacy Evaluable Set and mITT set.

At Baseline, Days 7, 14, 21, 30 and monthly thereafter, the study pulmonologist elicits the average daily amount of ventilation support used over the previous 1-2 weeks. As a secondary analysis, the amount of ventilation support will be categorized into None, >0-≤12 hours, >12-<16 hours, ≥16 or permanent ventilation. At each of the above time points, the count and percent of subjects in each ventilation support category will be displayed.

The average daily duration of ventilator support will be analyzed on the ITT set.

8.2. Secondary and Additional Efficacy Endpoint(s) and Analyses

8.2.1. Co-secondary Efficacy endpoint: Maintain ability to thrive

At baseline, a patient will be defined as not requiring parenteral nutrition if the child 1) did not use parenteral nutrition of any kind (i.e., gastrostomy, nasogastric tube, nasojejunal tube, etc.); and 2) demonstrated intact swallowing on baseline assessment such that the child did not receive a recommendation for implementation of parenteral nutrition prior to receipt of AVXS-101.

The proportion of patients who at baseline did not require parenteral nutrition (as defined) who maintain the ability to thrive at the time of the child's study visit when he or she reached 15 months of age will be summarized.

The ability to thrive is defined by meeting the following criteria:

- the ability to tolerate thin liquids as demonstrated through a formal swallowing test
- do not receive nutrition through mechanical support (i.e., feeding tube)
- maintains weight (> third percentile for age and gender as defined by WHO guidelines) at 15 months of age.

The observed proportion maintaining the ability to thrive will be compared to zero using a one-sided Exact Binomial Test. To make computation of the *p*-value possible, the value of 0.1% will be used in place of a literal zero.

This analysis will be performed on the "ability to thrive" ITT set.

8.2.2. Co-secondary Efficacy endpoint: Independence of ventilatory support

The proportion of patients who are independent of ventilatory support will be summarized. Independence of ventilatory support is defined as requiring no daily invasive or non-invasive ventilator support/usage at 15 months of age, in the absence of acute reversible illness and excluding perioperative ventilation. Ancillary devices used in pulmonary such as a cough assist device shall not be considered ventilatory support. For a child who experiences an acute reversible illness or requires perioperative ventilation at the time that he or she reaches 15

months of age, ventilatory independence will be defined as not requiring ventilatory support over the 2 week interval prior to the development of the acute reversible illness or perioperative use.

The observed proportion who are independent of ventilatory support will be compared to zero using a one-sided Exact Binomial Test. To make computation of the p-value possible, the value of 0.1% will be used in place of a literal zero.

These analyses will be performed on the Full Analysis Set and mITT set.

8.2.3. Efficacy Endpoint: Development of Significant Motor Function Milestones per Motor Milestone Development Survey

At Baseline, Day 30, Month 6, Month 12, and then every 3 months in Follow Up Year 2, the investigator is asked whether the subject currently exhibits evidence of achieving each of the motor milestones listed on the Motor Milestone Development Survey (see Appendix 21.1).

The number and percent of subjects who exhibit evidence of milestone achievement by the time of the Efficacy Data Cut-off will be summarized by cohort. The observed proportion attaining these milestones during the Observation Period will be compared to zero using a one-sided Exact Binomial Test. To make computation of the *p*-value possible, the value of 0.1% will be used in place of a literal zero.

The three additional motor milestones of the most interest in this population and age group are:

- Holds head up to vertical axis and legs extended
- Rolls over back to front
- Sits without support
  - as defined by WHO-MGRS guidelines – Child sits up straight with head erect for at least 10 seconds; child does not use arms or hands to balance body or support position
  - as defined by Nationwide Children's Hospital – sitting without support for 15 or more seconds

The number and percent of subjects who exhibit evidence of milestone achievement by the time of the Efficacy Data Cut-off will be summarized by cohort. The observed proportion attaining these milestones during the Observation Period will be compared to zero using a one-sided Exact Binomial Test. To make computation of the *p*-value possible, the value of 0.1% will be used in place of a literal zero.

As described above, independent of the investigator's assessment of milestones, an expert reviewer will evaluate videotapes of CHOP-INTEND and/or Bayley assessments conducted at clinic visits. The reviewer will judge whether the video reveals evidence of milestone achievement for each of the Motor Milestone Development Survey items. Parent(s)/legal guardian(s) may also share home videos demonstrating achievement of functional abilities and the reviewer may evaluate those videos for evidence of milestone achievement.

The number and percent of subjects, who through video evidence, exhibit milestone achievement of sitting without support; in addition to above two milestones, by the time of the Efficacy Data Cut-off, will be summarized at each visit by cohort.

These analyses will be performed on the Full Analysis Set and mITT set.

8.2.4. Efficacy Endpoint: Development of Significant Motor Function Milestones per Gross Motor Skills Checklist

At Baseline, Day 30, Month 6, Month 12, and then every 3 months in Follow-up Year 2, the investigator is asked whether the subject can currently perform motor skills listed on the Gross Motor Skills Checklist (see Appendix 21.3).

The number and percent of subjects who can perform each motor skill by the time of the Efficacy Data Cut-off will be summarized by cohort. The observed proportion performing each motor skill will be compared to zero using a one-sided Exact Binomial Test. To make computation of the *p*-value possible, the value of 0.1% will be used in place of a literal zero.

As described above, independent of the investigator's assessment of milestones, an expert reviewer will evaluate videotapes of CHOP-INTEND or Bayley assessments conducted at clinic visits. The reviewer will judge whether the video reveals performance of the motor skills in the Gross Motor Skills Checklist. Parent(s)/legal guardian(s) may also share home videos demonstrating achievement of functional abilities and the reviewer may evaluate those videos for evidence of motor skill performance.

The number and percent of subjects, through video evidence, who exhibit performance of each motor skill by the time of the Efficacy Data Cut-off will be summarized by cohort. The observed proportion performing each motor skill will be compared to zero using a one-sided Exact Binomial Test. To make computation of the *p*-value possible, the value of 0.1% will be used in place of a literal zero.

These analyses will be performed on the Full Analysis Set and mITT set.

8.2.5. Change from Baseline in CHOP-INTEND Score

Improvement in motor function will be quantified by change from baseline in CHOP-INTEND score. The CHOP-INTEND is a reliable, validated, 16-item assessment of motor function. Each item is scored 0 to 4 by a trained physical therapist during a comprehensive motor function assessment. For each applicable motor function item, motor function is assessed on both the left side and right side. A summary score is obtained by summing the individual best-side scores. The maximum score is 64, with a higher score representing higher clinical function. See the Missing Data section for special scoring instructions. CHOP-INTEND assessments will be conducted monthly except in certain cases, such as when a subject reaches a score ≥62 for ≥2 consecutive visits and the decision is made that the subject will complete only the Bayley assessment from that point forward.

The change from baseline will be analyzed by using mixed model with repeated measurement (MMRM). The model for full analysis set will include the change from baseline as the dependent variable, and fixed effects of cohort (or efficacy subgroup), visit, and a covariate of baseline, and interactions of cohort*visit, baseline*visit. The model for mITT set will be conducted with change from baseline as the dependent variable and fixed effects of visit and a covariate of baseline. An unstructured (general) covariance structure will be assumed initially to model the within-patient errors; however, if unstructured covariance results non-convergence, the variance component will be used. The least squares (LS) means, differences between LS

means, a 95% 2-sided CLs for each difference and the p-values from model effects will be reported for each scheduled visit.

8.2.6. Achievement of Selected Threshold Scores on CHOP-INTEND

The achievement of clinically relevant threshold scores will be presented.

- Ever reaching CHOP-INTEND score ≥40
- Ever reaching CHOP-INTEND score ≥50
- Ever reaching CHOP-INTEND score ≥60

The rationale for the clinical relevance of a score ≥40 is that this threshold is beyond that reported in the literature for maximum function amongst symptomatic patients with SMA Type 1 (Finkel et al 2014).

The rationale for the clinical relevance of a score ≥50 is that this approximates (slightly above the median) the range of scores for SMA Type 2 children reported in the PNCR study (Finkel et al 2014). Achieving this score would suggest the potential to gain milestones such as independent sitting associated with a Type 2 phenotype.

The rationale for the clinical relevance of a score ≥60 is that this value marks the effective ceiling for scoring clinical function using the CHOP-INTEND. The CHOP-INTEND includes items that are "age specific" and do not occur normally in healthy children over the age of around 6 months. For example, the grasp reflex changes to a volitional grasp so older children do not simply hold onto a finger when a “pull to sit” maneuver is attempted, but will instead let go when pulled. The same developmental impact applies for evaluating children held in horizontal suspension. This test was designed for infants with short life expectancy so it did not take developmental maturation into consideration. For this reason, it is expected that even healthy older children may not score above 60, rather than the full score of 64.

The number and percent of subjects who achieve each threshold by the time of the Efficacy Data Cut-off will be summarized by cohort. The observed proportion will be compared to zero using a one-sided Exact Binomial Test. To make computation of the *p*-value possible, the value of 0.1% will be used in place of a literal zero.

These analyses will be performed on the Full Analysis Set and mITT set.

8.2.7. Bayley Scales of Infant and Toddler Development©, version III

The maximum value on the CHOP-INTEND instrument is 64. During the course of the study some participants began achieving this maximum score, curtailing the ability to assess further incremental improvements in motor function. Therefore, the Bayley Scale of Infant and Toddler Development©, version III is co-administered by the physical therapist with the CHOP-INTEND once participants have reached a score of 60. The Bayley may be administered exclusively once participants have attained 2 consecutive CHOP-INTEND scores of ≥62, if agreed upon by AveXis and the investigator.

The Bayley Scales of Infant and Toddler Development (Bayley-III)© are recognized internationally as a comprehensive tool to assess children from as young as one month old. With Bayley-III©, it is possible to obtain detailed information from non-verbal children as to their

functioning. Children are assessed in the five key developmental domains of cognition, language, social-emotional, motor and adaptive behavior.

Bayley-III provides assessment of:

- Cognitive Development
- Receptive Communication
- Expressive Communication
- Fine Motor Development
- Gross Motor Development

Items are ordered by difficulty and should be administered in the order listed (with the exception of series items). The subtests were standardized by having examiners follow the item order provided. It also ensures that all pertinent items are administered (none are forgotten), and that reversal and discontinue rules are met quickly, with no extraneous items that may contribute to the fatigue of the child.

The Bayley may be administered exclusively once subjects have attained 2 consecutive CHOP-INTEND scores of 62. For purposes of this clinical trial, the following 5 Bayley Scale subtests will be administered: a.) Fine Motor Subtest; b.) Gross Motor Subtest; c.) Cognitive Subtest; d.) Language-Receptive Communication; and e.) Language-Expressive Communication. For each subtest, a raw score is computed on the basis of the number of items successfully completed. The raw score is then converted to a scaled score by the physical therapist according to the manual instructions.

A summary of changes from first recorded value in raw score performance on the Fine Motor Subtest and Gross Motor Subtests will be described by cohort or by the subgroup of study participants who are assessed using this motor function scale. These analyses will be performed on the Full Analysis Set and mITT set, as applicable.

The change from baseline of two motor subscales, i.e., fine motor development and gross motor development, will be analyzed by using mixed model with repeated measurement (MMRM) separately. The model for full analysis set will include the change from baseline as the dependent variable, and fixed effects of cohort (or efficacy subgroup), visit, and a covariate of baseline, and interactions of cohort*visit, baseline*visit. The model for mITT set will be conducted with change from baseline as the dependent variable and fixed effects of visit and a covariate of baseline. An unstructured (general) covariance structure will be assumed initially to model the within-patient errors; however, if unstructured covariance results non-convergence, the variance component will be used. The least squares (LS) means, differences between LS means, a 95% 2-sided CLs for each difference and the p-values from model effects will be reported for each scheduled visit.

8.2.8. ACTIVE-mini

ACTIVE-mini (Abilities Captured Through Interactive Video Evaluation – mini) was designed to measure spontaneous or elicited movements noninvasively using the Microsoft Kinect camera system. Through quantification of positional changes of hands and feet over time, ACTIVE-

mini has the ability to measure total movement volume and speeds of movement of each extremity separately.

The quantitative output of this assessment is still under development. Therefore, this endpoint is exploratory and will not be part of the 15 months of age Clinical Study Report. A separate analysis plan for scientific manuscripts and/or final Clinical Study Report may be written once the ACTIVE-mini quantification has been developed.

8.2.9. Motor Neuron Function assessed through Compound Motor Unit Potential (CMAP) and Motor Unit Number Estimation (MUNE)

The compound muscle action potential (CMAP) of a muscle or group of muscles is recorded with surface electrodes following supramaximal stimulation of the innervating nerve. The size of the CMAP can be assessed using the parameters of amplitude and area. The CMAP represents the summated depolarization of muscle fibers in a muscle or muscle group being recorded, and therefore gives a measure of total electrophysiological output. For the purpose of this trial, CMAP responses will be recorded from two nerve-muscle groups: tibialis anterior (peroneal) nerve and abductor digiti minimi (ulnar) nerve. Amplitude and area will be recorded for each CMAP measurement. Amplitude is recorded in millivolts (mV) and area in millivolt seconds (mVsec).

Motor Unit Number Estimation (MUNE) allows estimation of the number of functional motor units innervating a muscle or group of muscles. To perform MUNE of the abductor digiti minimi muscle, up to 10 single motor unit potentials will be recorded and averaged. MUNE is calculated as:

MUNE = (CMAP amplitude in microvolts) / mean (SMUP), where SMUP <25 microvolts are excluded.

MUNE is calculated for the Abductor digiti mini muscle only.

Detailed procedures for measuring CMAP and MUNE are included in Appendix 21.4.

A summary of changes from first recorded value in CMAP responses from tibialis anterior-peroneal nerve (Peroneal CMAP), abductor digiti minimi-ulnar nerve (Ulnar CMAP) and MUNE as well will be described by cohort or by efficacy subgroup. These analyses will be performed on the Full Analysis Set and mITT set, as applicable.

The change from baseline of these two CMAPs will be analyzed by using mixed model with repeated measurement (MMRM) separately. The model for full analysis set will include the change from baseline as the dependent variable, and fixed effects of cohort (or efficacy subgroup), visit, and a covariate of baseline, and interactions of cohort*visit, baseline*visit. The model for mITT set will be conducted with change from baseline as the dependent variable and fixed effects of visit and a covariate of baseline. An unstructured (general) covariance structure will be assumed initially to model the within-patient errors; however, if unstructured covariance results non-convergence, the variance component will be used. The least squares (LS) means, differences between LS means, a 95% 2-sided CLs for each difference and the p-values from model effects will be reported for each scheduled visit.

8.2.10. Pathological status of muscles quantified using Electrical Impedance Myography

These data are exploratory and will not be part of the 15months of age Clinical Study Report. A separate analysis plan for scientific manuscripts and/or final Clinical Study Report may be written in the future.

9. ANALYSIS OF PHARMACOKINETICS

Not applicable to this SAP.

10. SAFETY

10.1. Extent of Exposure

The actual, weight-adjusted dose, in vg/kg, of AVXS-101 administered during the infusion will be summarized by cohort, as well as the site of infusion, duration of infusion, whether the entire volume was delivered, and whether the infusion was interrupted.

10.2. Treatment Compliance

At protocol-specified baseline, the actual, weight-adjusted dose, in vg/kg, of AVXS-101 administered during the infusion was recorded as complete for all 15 patients. Volume and dosing compliance of AVXS-101 will be summarized according to cohort.

10.3. Adverse Events / Adverse Drug Reactions

10.3.1. Treatment-Emergent Adverse Events

Adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). Treatment-emergent adverse events are defined as any event that begins or worsens in severity after initiation of study drug through 30 days after last study visit. Events where the onset date is the same as the study drug start date are assumed to be treatment-emergent. If an incomplete onset date was collected for an adverse event, the event will be assumed to be treatment-emergent, unless there is other evidence that confirms that the event was not treatment-emergent (e.g., the event end date was prior to the study drug start date).

10.3.2. Tabulations of Treatment-Emergent Adverse Events

Adverse event data will be summarized and presented using primary MedDRA system organ classes (SOCs) and preferred terms (PTs) according to the version of the MedDRA coding dictionary used for the study at the time of database lock. The actual version of the MedDRA coding dictionary used will be noted in the clinical study report. The system organ classes will be presented in alphabetical order and the preferred terms will be presented in alphabetical order within each system organ class.

Adverse events will be presented by cohort and overall.

10.3.2.1. Adverse Event Overview

An overview of adverse events will be presented for each treatment group consisting of the number and percentage of subjects experiencing at least one event for the following adverse event categories:

- Any treatment-emergent adverse event;
- Treatment-emergent adverse events with a "possibly related", "probably related", "definitely related" of being related to AVXS-101
- Severe treatment-emergent adverse events;
- Serious treatment-emergent adverse events;

- Treatment-emergent adverse events leading to discontinuation of patient from study;
- Treatment-emergent adverse events leading to death;
- Deaths.

For each adverse event presented in the overview by cohort.

10.3.2.2. Adverse Event by SOC and PT

The following summaries of adverse events will be generated:

- Treatment-emergent adverse events;
- Treatment-emergent adverse events with a "possibly related" of being related to AVXS-101;
- Treatment-emergent adverse events with a "probably related" of being related to AVXS-101;
- Treatment-emergent adverse events with a "definitely related" of being related to AVXS-101;
- Serious treatment-emergent adverse events;
- Moderate or severe treatment-emergent adverse events;
- Severe treatment-emergent adverse events;
- Grade 3 or 4 (see definition below) treatment-emergent adverse events;
- Treatment-emergent adverse events leading to discontinuation of patient from study;
- Treatment-emergent adverse events leading to death;
- Treatment-emergent adverse events leading to concomitant medication use (events with other action taken of "concomitant medication prescribed").

For all adverse event summaries, the number and percentage of subjects experiencing treatment-emergent adverse events will be tabulated according to SOC and PT for each treatment group. Subjects reporting more than one adverse event for a given PT will be counted only once for that term (most severe incident for the severity tables and most related incident for the relationship tables). Subjects reporting more than one adverse event within a SOC will be counted only once for that SOC. Subjects reporting more than one adverse event will be counted only once in the overall total.

The percentage of subjects experiencing treatment-emergent adverse events, treatment emergent adverse events with a "possibly related", "probably related", or "definitely related" of being related to study drug, moderate or severe treatment-emergent adverse events, and severe treatment-emergent adverse events will be compared between cohorts using Fisher's exact tests. Only P values ≤ 0.100 when rounded to three digits will be presented.

A listing by treatment group of treatment-emergent adverse events grouped by body system and preferred term with subject numbers will be created.

10.3.2.3. Adverse Event by PT

The number and percentage of subject experiencing treatment-emergent adverse events will be tabulated according to preferred term and sorted by overall frequency in the total of both cohorts combined. Similar summaries will be provided for moderate to severe treatment-emergent adverse events and treatment-emergent adverse events with a "possibly related" of being related to AVXS-101. Percentages will be compared between treatment groups using Fisher's exact tests. Only P values ≤0.100 when rounded to three digits will be presented.

10.3.2.4. Adverse Events of Special Interest

The following specific treatment-emergent adverse events of special interest, which may be searched using Standardized MedDRA queries, will be summarized:

Elevated liver enzymes

For this adverse event of interest, the number and percentage of subjects experiencing at least one treatment-emergent adverse event in the search for the event of interest will be presented for each cohort overall.

10.3.2.5. Adverse Events by Maximum Severity

Treatment-emergent adverse events and treatment-emergent adverse events with a "possibly related", "probably related", or "definitely related" of being related to AVXS-101 will be summarized by maximum severity of each preferred term. If a subject has an adverse event with unknown severity, then the subject will be counted in the severity category of "unknown," even if the subject has another occurrence of the same event with a severity present.

10.3.2.6. Adverse Events by Maximum Severity Grade Level

Treatment-emergent adverse events will be summarized by maximum severity grade level of each preferred term. Each preferred term will be assigned to a grade level based on severity and seriousness, adapted from CTCAE Version 4.03 for grading severity of adverse events.

All serious adverse events will be categorized according to Grade definition.

If a subject has a non-serious adverse event with unknown severity, then the subject will be counted in the severity grade level category of "unknown," even if the subject has another occurrence of the same event with a severity present.

10.3.2.7. Adverse Events by Maximum Relationship

Treatment-emergent adverse events also will be summarized by maximum relationship of each preferred term to study drug (AVXS-101), as assessed by the investigator. If a subject has an adverse event with unknown relationship, then the subject will be counted in the relationship category of "unknown," even if the subject has another occurrence of the same event with a relationship present. The only exception is if the subject has another occurrence of the same adverse event with a relationship assessment of "possibly related." In this case, the subject will be counted under the "possibly related" category.

10.4. Laboratory Evaluations

10.4.1. Analysis of Laboratory Data

Data collected from OpenClinica and Rave, including additional laboratory testing due to any SAEs, will be used in all analysis.

10.4.2. Variables and Criteria Defining Abnormality

Hematology variables include: hematocrit, Hemoglobin, red blood cell (RBC) count, white blood cell (WBC) count, lymphocytes, neutrophils, lymphocytes, monocytes, eosinophils, basophils, bands, Mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH), mean cell hemoglobin concentration (MCHC), red cell distribution width (RDW), prothrombin time (PT), partial thromboplastin time (PTT), and international normalized ratio (INR).

Chemistry variables include: albumin, alanine aminotransferase (ALT/SGPT), alkaline phosphatase, amylase, aspartate aminotransferase (AST/SGOT), blood urea nitrogen (BUN), carbon dioxide (CO2), chloride, creatinine, gamma glutamyl transferase (GGT), glucose, serum total bilirubin, direct bilirubin, (CK), (CK-MB), potassium, sodium, troponin.

Urinalysis variables include: specific gravity, pH, ketones, glucose, protein, blood, leukocyte esterase, nitrites, bilirubin, urobilirubin, red blood cell (RBC) count, white blood cell (WBC) count, yeast, squamous epithelial cells, casts, crystals, bacteria.

Immunology variables include: mother's serum binding antibody titer to AAV9, serum binding antibody titer to AAV9.

The Criteria for Potentially Clinically Significant Values will be based upon CTCAE Version 4.3 criteria for Grade 2 or higher adverse event unless otherwise specified. Laboratory Reference Ranges provided by Nationwide Children's Hospital, Columbus OH.

The Criteria for Potentially Clinically Significant (PCS) Laboratory Findings are described in Table 8 and Table 9.

Table 8 Criteria for Potentially Clinically Significant Hematology Values

Test / Units	Very Low (VL)	Very High (VH)
Hemoglobin (g/dL)	<7.5 g/dl	>15.5 g/dl
Platelets Count (k cells/mm^3)	<75,000/mm^3	
White Blood Cell Count (k cells/mm^3)	<3.0/mm^3	
Neutrophil (%)	<18%	
Lymphocyte (cell/mm^3)	<5000	
Eosinophil (%)		>5%
aPTT		>1.5 x ULN
International Normalized Ratio		>1.5 x ULN

Table 9 Criteria for Potentially Clinically Significant Chemistry Values

Test / Units	Very Low (VL)	Very High (VH)
ALT/SGPT		>3.0 x ULN
AST/SGOT		>3.0 x ULN
Alkaline Phosphatase		>2.5 x ULN
Total Bilirubin (mg/dL)		>1.5 x ULN
Creatinine		>1.5 x ULN
BUN		>1.5 x ULN
Serum CK		>2.5 x ULN
Sodium (mmol/L)	<130 mmol/L	>150 mmol/L
Troponin I		>0.05 ng/ml
Potassium (mmol/L)	<2.5 mmol/L	>5.5 mmol/L
Glucose	<55 mg/dl	>160 mg/dl
Albumin	<30 g/L	
GGT		>2.5 x ULN
Amylase		>1.5 x ULN
CO2	<16 mmol/L	

10.4.3. Statistical Methods

Clinical laboratory test will be summarized by cohort at each visit during the treatment period. The baseline value will be the last measurement on or before the day of the infusion of study drug. This same baseline value will be used for all changes from baseline tables in the treatment period and post-treatment period.

Mean changes from baseline to each post-baseline visit, including applicable post-treatment visits, will be summarized for each protocol-specified laboratory parameter with the baseline mean, visit mean, change from baseline mean, standard deviation, minimum, maximum, and median.

During the treatment period, laboratory data values will be categorized as low, normal, or high based on normal ranges of the laboratory used in this study. Shift tables from baseline to minimum value, maximum value, and final values during the treatment period will be created. The shift tables will cross tabulate the frequency of patients with baseline values below/within/above the normal range versus minimum/maximum/final values below/within/above the normal range.

The number and percentage of patients with post-baseline values during the treatment period meeting the specified criteria for Potentially Clinically Significant (PCS) laboratory values (defined in Table 8 and Table 9) will be summarized by cohort. A post-baseline value must be

more extreme than the baseline value to be considered a PCS finding. A listing will be provided that presents all of the lab values for the patients meeting PCS criteria during treatment.

For hemoglobin and the liver function tests (LFTs) of ALT, AST, alkaline phosphatase, and total bilirubin, the number and percentage of patients in each treatment group with a maximum CTCAE Grade of 1, 2, 3, or 4 (see definitions in Table 10) at any post-baseline visit (regardless of the baseline value) through the end of treatment (i.e., Final Treatment Value) will be summarized. All LFT tables will include summary rows for the number and percentage of subjects with at least Grade 2 and at least Grade 3 laboratory abnormalities. The hemoglobin table will include a summary row for the number and percentage of patients with at least a Grade 2 laboratory abnormality. Cohort comparisons of the percentage of subjects experiencing a value meeting at least Grade 2 and at least Grade 3 (as reported in the summary row(s)) will be performed using Fisher's exact tests. Only P values ≤0.100 when rounded to three digits will be presented. Accompanying listings of all ALT, AST, total, indirect and direct bilirubin, and alkaline phosphatase will be created for any patients who had at least a Grade 3 ALT, AST, alkaline phosphatase, or total bilirubin. A listing of hematology results will be provided for subjects with hemoglobin abnormalities.

For subjects with a Grade 3 or higher total bilirubin elevation, a listing of treatment-emergent adverse events (defined as preferred terms within the "Cholestasis and jaundice of hepatic origin" (broad search) SMQ, excluding preferred terms within the "Investigations" SOC) will be provided.

Table 10 Definitions of CTCAE Grades 1, 2, 3, and 4

Test	Grade 1	Grade 2	Grade 3	Grade 4
ALT/SGPT	>ULN – 3 × ULN	>3 – 5 × ULN	>5 – 20 × ULN	>20 × ULN
AST/SGOT	>ULN – 3 × ULN	>3 – 5 × ULN	>5 – 20 × ULN	>20 × ULN
Alkaline phosphatase	>ULN – 2.5 × ULN	>2.5 – 5 × ULN	>5 – 20 × ULN	>20 × ULN
Total bilirubin	>ULN – 1.5 × ULN	>1.5 – 3 × ULN	>3 – 10 × ULN	>10 × ULN
Hemoglobin decreased	<LLN – 8.5 g/dL	<8.5 – 7.5 g/dL	<7.5 – 6.5 g/dL	<6.5 g/dL

Reference Ranges, Kathleen Nicol, Department of Pathology and Laboratory Medicine of Nationwide Children's Hospital, Columbus, Ohio, 1-11.

The number and percentage of subjects in each cohort meeting the following criteria will be summarized for each treatment period:

- ALT ≥3 × ULN and total bilirubin value ≥2 × ULN
- ALT ≥3 × ULN and total bilirubin value <2 × ULN;
- ALT >5 × ULN (equivalent to Grade 3 or higher) and total bilirubin value <2 × ULN;
- ALT <3 × ULN and total bilirubin ≥2 × ULN.

A patient or event will be counted if the post-baseline laboratory values meet the above criteria regardless of the baseline laboratory value (i.e., the post-baseline laboratory value does not need to be worse than the baseline laboratory value). The maximum ratio relative to the ULN will be used to determine if patients meet the criteria listed above. For patients meeting the ALT ≥3 × ULN and total bilirubin value ≥2 × ULN criterion during the Treatment Periods, a corresponding listing of all ALT, AST, alkaline phosphatase, and total, direct, and indirect bilirubin values will be provided.

For subjects meeting the criterion of ALT <3 × ULN and total bilirubin ≥2 × ULN, the number and percentage of subjects with a total bilirubin value in the categories of ≤ULN, >ULN – <2 × ULN, and ≥2 × ULN at the Final Treatment Visit will be summarized.

In addition, for patients meeting the criterion of ALT <3 × ULN and total bilirubin ≥2 × ULN (based on the maximum ratio relative to the ULN), the ratio of indirect bilirubin to total bilirubin will be calculated. The following summary statistics will be presented for each treatment group for the ratio at baseline and for the ratio associated with the peak total bilirubin value during the Treatment Period: sample size, mean, standard deviation, minimum, maximum, and median. In addition, the number and percentage of patients with a ratio <0.75 and <0.50 will be presented for baseline and peak.

10.4.4. Drug-Induced Liver Injury

Drug-induced liver injury (DILI) has been the most frequent single cause of safety-related drug marketing withdrawals and as such with this pediatric population (SMA patients) and unique intervention (gene therapy), Sponsor took care to assess hepatic test results as drugs can cause liver injuries by many different mechanisms. Severe DILI cases rarely have been seen in drug development programs of significantly hepatotoxic drugs that do cause such injury. Evidence of hepatocellular injury is thus a necessary, but not sufficient, signal of the potential to cause severe DILI (note, however, that the drugs causing hepatic injury through mitochondrial toxicity may not cause early hepatotoxicity).

It is possible that although a drug may not cause severe liver injury, it could still result in laboratory evidence of mild, transient hepatic injury, with leakage of liver enzymes and the appearance in serum of elevations in aminotransferase activities to levels of 3, 5, and sometimes greater than 5 times the upper limits of normal (ULN). The liver enzyme data was evaluated according to these criteria and according to Hy's Law.

A finding of ALT elevation, usually substantial, seen concurrently with bilirubin >2xULN, identifies a drug likely to cause severe DILI (fatal or requiring transplant) at a rate roughly 1/10 the rate of Hy's Law cases. It is critical to rule out other causes of injury (e.g., other drugs or viral hepatitis) and to rule out an obstructive basis for the elevated bilirubin, so that alkaline phosphatase (ALP) should not be substantially elevated. In all cases to date, the small number of Hy's Law cases has arisen on a background of an increased incidence of more modest signs of hepatocellular injury (e.g., greater incidence of 3xULN elevations in AT than seen in a control group).

Briefly, Hy's Law cases have the following three components:

- The drug causes hepatocellular injury, generally shown by a higher incidence of 3-fold or greater elevations above the ULN of ALT or AST than the (nonhepatotoxic) control drug or placebo.
- Among trial subjects showing such AT elevations, often with ATs much greater than 3xULN, one or more also show elevation of serum TBL to >2xULN, without initial findings of cholestasis (elevated serum ALP).
- No other reason can be found to explain the combination of increased AT and TBL, such as viral hepatitis A, B, or C; preexisting or acute liver disease; or another drug capable of causing the observed injury.

For patients enrolled in this study these criteria are assessed in order to determine both general liver function enzyme changes (LFEs) according to change from baseline and over course of study. In addition, the assessment of changes in ALT or AST relative to TBL allowed assessment if a signal related to DILI occurred in any individual or group of patients.

10.5. Vital Signs

Vital signs (pulse, respiration, temperature, diastolic blood pressure, body weight, systolic blood pressure, pulse oximetry) will be examined at each visit. Clinically-significant, treatment-emergent findings will be reported as adverse events.

A summary of changes from first recorded value in vital signs will be described at each visit by cohort on the safety set. In addition, vital signs results will be flagged as Potentially Clinically Significant (PCS) if they meet the pre-specified criteria outlined in Table 6. The number and percent of subjects meeting each PCS criterion will be summarized starting at Day 1 and continuing through Month 24.

The Criteria for Potentially Clinically Significant Vital Sign Findings are presented in Table 11.

Table 11 Criteria for Potentially Clinically Significant Vital Sign Values

Test / Measurement	Very Low (VL)	Very High (VH)
Systolic blood pressure*	<67	>110
Diastolic blood pressure*		>64
Pulse rate**	Below 5th percentile for age	Above 95th percentile for age
Body Weight***	Weight <3rd percentile for age and gender	Weight>97th percentile for age and gender
Temperature	<35°C	>39°C

* Adapted from NHLBI Reference Ranges for age

** Bonafide et al. Pediatrics 2013;131:e1150–e115

*** Based upon NHANES III Percentiles for age and gender

10.6. ECG

10.6.1. Additional Safety Endpoint: Cardiovascular Safety Evaluations

Echocardiograms and electrocardiograms will be conducted at baseline, Day 30, every 6 months to Month 12, then every 3 months to Month 24. Clinically significant, treatment-emergent findings will be reported as adverse events.

In addition, selected ECG results will be flagged as Potentially Clinically Significant if they meet the pre-specified criteria outlined in Table 7. The number and percent of subjects meeting each PCS criterion will be summarized starting at Day 30 and continuing through Month 24.

Table 12 Criteria for Potentially Clinically Significant EKG Values

Panel	Parameter	Low/High	PCS Criteria
ECG[1]	Heart Rate	Low	Below 5th percentile for age*
	Heart Rate	High	Above 5th percentile for age*
	Heart Rate	Change	Inc > = 20 bpm
	QTcB	High	≥440 msec
	QTcB	Change	Inc > = 30 msec
	QTcF	High	≥440 msec
	QTcF	Change	Inc > = 30 msec

*Bonafide et al. Pediatrics 2013;131:e1150–e115

10.7. Physical Examination

Treatment-emergent abnormal findings on physical exam will be tracked as adverse events. Any post-infusion abnormal physical exam findings will be listed by subject with the corresponding result on the baseline physical exam.

10.8. Other Safety

10.8.1. Additional Safety Endpoint: Use of Non-oral Feeding Support

The number and percent of subjects who ever used non-oral feeding support through to the date when all patients reach 15 months of age will be summarized by cohort, with the type of feeding tube (gastrostomy with Nissen fundoplication, gastrostomy without Nissen fundoplication, nasogastric, or nasojejunal) categorized.

Among those subjects who never used feeding support prior to enrollment, the age at first use of non-oral feeding support will be estimated by Kaplan-Meier analysis for the Efficacy Data cut-off observation period. Ninety-five percent confidence intervals will be computed for the 25th percentile, median, and 75th percentiles of the survival function.

These analyses will be conducted for each cohort.

10.8.2. Additional Safety Endpoint: Immunologic Response

T-cell immunoreactivity to AAV9 and hSMN will be monitored by the collection of samples at baseline, Day 7, Day 14, Day 21, Day 30, Day 60, Day 90 and every 3 months thereafter. Antibody titer levels are measured through ELISA immunoassay. Antibody titers >1:50 are considered positive for antibody response while antibody titers ≤1:50 are considered negative. The number and percent of subjects responding Positive or Negative for antibody response at each time point will be summarized by cohort. Furthermore, the distribution of subjects by titer level will be summarized by cohort.

Secondly, T-cell response to AAV9 and SMN will be measured continuously through the quantification of number of Spot Forming Cells (SFC) per million Peripheral Blood Mononuclear Cells (10^6PBMC). T-cell response to AAV9 will be measured separately in 2 peptide pools. T-cell response to SMN will be measured in a single peptide pool.

The number of SFC/10^6PBMC will be summarized at each sample time point. For each post-infusion time point, change from baseline and percent change from baseline will be computed and tested against zero using a one-sample Student t-test and Wilcoxon Signed Rank test.

Summaries will be presented separately for Cohort 1 and Cohort 2.

Subject E.01 received gene therapy infusion prior to the protocol amendment that stipulated prophylactic administration of prednisolone prior to gene therapy. Without such prophylactic therapy to dampen the host immune response to the AAV based therapy, Subject E.01's T-cell response measured by ELiSPOT could have been systematically higher than those subjects who did receive the prophylactic prednisolone and thereby skew the overall summary results higher.

Medical coding will be performed on the verbatim terms from the Concomitant Medications eCRF form and the Prednisolone eCRF form using the WHO Drug version March, 2016 C Final 2 coding dictionary. The version of the coding dictionary will undergo update on an annual basis (March update) for this study.

Therefore, a secondary analysis of T-cell response to AAV9 and SMN will be performed on the subset of Safety Analysis Set subjects excluding E.01.

10.8.3. Additional Safety Endpoint: Pulse Oximetry

Pulse oximetry will be measured throughout the study through a small infrared light attached to the end of the patient's finger.

11. HEALTH ECONOMICS

Not applicable for this SAP.

12. INTERIM ANALYSES

Refer to Table 1.

13. CHANGE FROM ANALYSIS PLANNED IN PROTOCOL

See Amended Protocol AVXS-101-CL-101 Version 15.0 dated 14-Nov-2016 for changes from analysis planned in protocol.

14. REFERENCE LIST

1 Finkel RS, McDermott MP, Kaufmann P, Darras BT, Chung WK, Sproule DM, Kang PB, Foley AR, Yang ML, Martens WB, Oskoui M, Glanzman AM, Flickinger J, Montes J Dunaway S, O'Hagen J, Quigley J, Riley S, Benton M, Ryan PA, Montgomery M, Marra J, Gooch C, De Vivo DC. Observational study of spinal muscular atrophy type I and implications for clinical trials. Neurology. 2014 Aug 26;83(9):810-7.

2 Cancer Therapy Evaluation Program, Common Terminology Criteria for Adverse Events, Version 4.3, DCTD, NCI, NIH, DHHS (http://ctep.cancer.gov), Publish Date: May 28, 2009.

3 Bonafide CP, Brady PW, Keren R, Conway PH, Marsolo K, Daymont C. Development of Heart and Respiratory Rate Percentile Curves for Hospitalized Children. Pediatrics. 2013 April; 131(4): e1150–e1157.

4 Reference Ranges, Kathleen Nicol, Department of Pathology and Laboratory Medicine of Nationwide Children's Hospital, Columbus, Ohio, 1-11.

5 Wijnhoven TM, de Onis M, Onyango AW, et al. Assessment of gross motor development in the WHO Multicentre Growth Reference Study. *Food Nutr Bull*. 2004;25(1 Suppl): S37-45.

15. PROGRAMMING CONSIDERATIONS

See TLF documentation for details.

16. QUALITY CONTROL

Refer to Quality Control Plan for SAS programs.

17. INDEX OF TABLES

17.1. LIST OF TABLES

Table No.	Title
Table xx.x	Summary of Treatment-Emergent Adverse Events through 15 months of age
Table xx	Incidence of Adverse Events through 15 Months of Age by System Organ Class and Preferred Term
Table xx	Summary of Treatment-Emergent Adverse Events through Efficacy Data cut-off
Table xx	Incidence of Adverse Events through Efficacy Data cut-off by System Organ Class and Preferred Term
Table xx	Incidence of Serious Adverse Events through Efficacy Data cut-off by System Organ Class and Preferred Term
Table xx	Distribution of Subjects by Use of Feeding Support through 15 months of age
Table xx	Time of First Use of Non-oral Feeding Support through Efficacy Data cut-off
Table xx	Distribution of Subjects by Serum-Binding Titer Values to AAV9 through 15 months of age
Table xx	Distribution of Subjects by Serum-Binding Titer Values to hSMN through 15 months of age
Table xx	Immunoreactivity to AAV9: Number of Spot Forming Cells per Million Peripheral Blood Mononuclear Cells(106PBMC)
Table xx	Immunoreactivity to AAV9: Number of Spot Forming Cells per Million Peripheral Blood Mononuclear Cells(106PBMC) Subjects Treated with Prophylactic Prednisolone
Table xx	Immunoreactivity to hSMN: Number of Spot Forming Cells per Million Peripheral Blood Mononuclear Cells(106PBMC)
Table xx	Immunoreactivity to hSMN: Number of Spot Forming Cells per Million Peripheral Blood Mononuclear Cells(106PBMC) Subjects Treated with Prophylactic Prednisolone
Table xx	Summary of Potentially Clinically Significant Lab Findings- Blood Chemistry
Table xx	Summary of Potentially Clinically Significant Lab Findings- Hematology
Table xx	Summary of Potentially Clinically Significant Lab Findings- Coagulation
Table xx	Summary of Potentially Clinically Significant ECG Findings
Table xx	Summary of Potentially Clinically Significant Vital Signs Findings

18. INDEX OF FIGURES

18.1. LIST OF FIGURES

19. INDEX OF LISTINGS

Index to be added once the TLF shells are generated.

20. MOCK-UPS

Refer to TLF Mock-ups.

21. APPENDICES

21.1. Motor Milestone Development Survey

4 months	Prone: Holds head up to vertical axis & legs extended
6 months	Supine: Rolls over back to front
9 months	Sits alone, with back straight
12 months	Cruises, holding on and may stand without help
15 months	Walks independently
18 months	Runs, walks down stairs one hand held
24 months	Up and down stairs, one step at a time; jumps both feet off floor
30 months	Reciprocal stair climbing; stands on one foot
36 months	36 months Reciprocal stairs going down; rides tricycle
48 months	Hops on one foot; throws ball overhand
60 months	Able to skip

21.2. CHOP-INTEND

CHILDREN'S HOSPITAL *of* PHILADELPHIA INFANT TEST OF NEUROMUSCULAR DISORDERS

Time of evaluation: _____ **(AM/PM)** **Hours off BiPAP at testing:** _____ **(h)**

Item	Position	Test Procedure	Graded Response		Score	
1 Spontaneous movement (Upper extremity)	Supine	Observe throughout testing May unweight limb or stimulate infant to facilitate response	Antigravity shoulder movement (achieves elbow off surface)	4	L	Best side:
			Antigravity elbow movement (achieves hand and forearm off surface)	3		
			Wrist movement	2		
			Finger movement	1	R	State:
			No movement of limbs	0		
2 Spontaneous movement (Lower extremity)	Supine	Observe throughout testing May unweight limb or stimulate infant to facilitate response	Antigravity hip movement (achieves feet and knees off surface)	4	L	Best side:
			Antigravity hip adduction/internal rotation (knees off surface)	3		
			Active gravity eliminated knee movement	2		
			Ankle movement	1		
			No movement of limbs	0	R	State:
3 Hand grip	Supine	Grip strength: place finger in palm and lift until shoulder comes off surface observe when infant loses grasp May use toy of similar diameter for older children	Maintains hand grip with shoulder off bed	4	L	Best side:
			Maintains grip with elbow off surface (shoulders on surface)	3		
			Maintains grip with forearm off surface (elbow supported on surface)	2		
			Maintains grip only with no traction	1	R	State:
			No attempt to maintain grasp	0		
4 Head in midline with visual stimulation*	Supine head midline	Visual stimulation is given with toy. ***If head is maintained in midline for 5 seconds:*** Place head in maximum available rotation and provide visual stimulation to encourage midline	Rotates from maximum rotation to midline	4	L>R	Best side:
			Turns head part way back to midline	3		
			Maintains midline for 5 or more seconds	2		
			Maintains midline, less than 5 seconds	1		
			Head falls to side, no attempts to regain midline	0	R>L	State:
5 Hip adductors	Supine, no diaper	Hips flexed and adducted Feet hip width apart and thighs parallel, knees slightly apart	Keeps knee off surface of bed >5 sec or lifts foot off surface	4	L	Best side:
			Keeps knees off surface of bed 1-5 sec	2		
			No attempt to maintain knees off surface	0	R	State:

Time of evaluation: _____ **(AM/PM)** **Hours off BiPAP at testing:** _____ **(h)**

Item	Position	Test Procedure	Graded Response		Score	
6 Rolling: elicited from legs*	Supine (arms at side) Keep side tested up roll away from the Side tested	**1.** Holding infant's lower thigh, flex hip and knee and adduct across midline bringing pelvis vertical maintain traction and ***pause in this position***. **2.** If infant rolls to side apply traction at a 45º diagonal to body and pause to allow infant to attempt to de-rotate body	When traction is applied at the end of the maneuver, rolls to prone with lateral head righting	4	To R	Best side:
			Rolls through side lying into prone without lateral head righting, clears weight-bearing arm to complete roll	3		
			Pelvis, trunk and arm lift from support surface, head turns and rolls onto side, arm comes thru to front of body	2		
			Pelvis and trunk lift from support surface and head turns to side. Arm remains behind trunk	1	To L	State:
			Pelvis lifted passively off support surface.	0		
7 Rolling: elicited from arms*	Supine (arms at side) Keep side tested up roll away from the Side tested	**1.** Hold infant at the elbow move toward opposite shoulder maintain traction on limb and ***pause with the shoulders vertical*** allow infant to de-rotate **2.** if the pelvis achieves vertical continue to provide traction	Rolls to prone with lateral head righting	4	To R	Best side:
			Rolls into prone without lateral head righting; must clear weight-bearing arm completely to finish roll	3		
			Rolls onto side, leg comes thru and adducts, bringing the pelvis vertical	2		
			Head turns to side and shoulder and trunk lift from surface	1	To L	State:
			Head turns to side; body remains limp or shoulder lifts passively	0		
8 Shoulder and elbow flexion And horizontal abduction	Side-lying with upper arm at 30º of shoulder extension and elbow flexion and supported on body (restrain lower arm if needed)	Prompt reach for a toy presented at arm's length at shoulder level (may provide stimulation and ***observe spontaneous movement***)	Clears hand from surface with antigravity arm movement	4	L	Best side:
			Able to flex shoulder to 45 degrees, without antigravity arm movement	3		
			Flexes elbow after arm comes off body	2		
			Able to get arm off body	1		
			No attempt	0	R	State:
9 Shoulder flexion & Elbow flexion	Sitting in lap or on mat with head and trunk support (20º recline)	Present stimulus at midline and at shoulder level at arm's length (may provide stimulation and ***observe spontaneous movement***)	Abducts or flexes shoulder to 60 degrees	4	L	Best side:
			Abducts or flexes shoulder to 30 degrees	3		
			Any shoulder flexion or abduction	2		
			Flexes elbow only	1		
			No attempt to lift arm	0	R	State:
10 Knee extension	Sitting in lap or over edge of mat with head and trunk support (20º recline) thigh horizontal to ground	Tickle plantar surface of foot or gently pinch toe	Extends knee to >45 degrees	4	L	Best side:
			Extends knee 15 to 45 degrees	2		
			Any visible knee extension	1		
			No visible knee extension	0	R	State:

Time of evaluation: _____ (AM/PM) **Hours off BiPAP at testing: _____ (h)**

Item	Position	Test Procedure	Graded Response		Score	
11 Hip flexion and foot dorsiflexion	Hold infant against your body with legs free, facing outward. Support at the abdomen with the child's head resting between your arm and thorax	Stroke the foot or pinch the toe	Hip flexion or knee flexion >30º	4	L	Best side:
			Any hip flexion or knee flexion	3		
			Ankle dorsiflexion only	2		
			No active hip, knee or ankle motion	1	R	State:
			Hip flexion or knee flexion >30º	0		
12 Head control*	Sitting with support at the shoulders and trunk erect	Place the infant in ring sit with head erect and assistance given at the shoulders (front and back) (***may delay scoring a grade of 1 and 4 until end of test***).	Attains head upright from flexion and turns head side to side	4	L	Best side:
			Maintains head upright for >15 sec (for bobbing head control score a 2)	3		
			Maintains head in midline for >5 sec. with the head tipped in up to 30º of forward flexion or extension	2		
			Actively lifts or rotates head twice from flexion within 15 seconds (do not credit if movement is in time with breathing)	1	R	State:
			No response, head hangs	0		
13 Elbow flexion Score with item 14	Supine	**Traction response:** pull to sit extend arms at 45 degree angle, to point of nearly lifting head off surface	Flexes elbow	4	L	Best side:
			Visible biceps contraction without elbow flexion	2		
			No visible contraction	0	R	State:
14 Neck Flexion Score with item 13	Supine	**Traction response:** hold in neutral proximal to wrist and shoulder at 45º, ***to point of nearly lifting head off surface***	Lifts head off bed	4		Score:
			Visible muscle contraction of SCM	2		
			No muscle contraction	0		State:
15 Head/Neck Extension (Landau)	Ventral suspension: Prone, held in one hand upper abdomen	Stoke along spine from neck to sacrum. The coronal axis of the head when parallel to the bed surface = 0 degrees (horizontal)	Extends head ***to*** horizontal plane or above	4		Score:
			Extends head partially, but not to horizontal	2		
			No head extension	0		State:
16 Spinal Incurvation (Galant)	Ventral suspension: Prone, held in one hand upper abdomen	Stroke Right then Left throacolumbar paraspinals or tickle abdomen or foot or tilt in infants with For infant over 10 kg knees and head may touch	Twists pelvis towards stimulus off axis	4	L	Best side:
			Visible paraspinal muscle contraction	2		
			No response	0	R	State:

Time of evaluation: _____ **(AM/PM)** **Hours off BiPAP at testing:** _____ **(h)**

Item	Position	Test Procedure	Graded Response		Score	
Total score, best score on each side for each item (maximum 64 points):						
Comments:						

* Adapted from the Test of Infant Motor Performance, Campbell, SK; et al. 2001.

Contractures:

□ None

L □ R □ Knee flexion

L □ R □ Ankle plantar flexion
(Present <20 degrees knee extended)

L □ R □ Hip adductor L □ R □ ITB contracture

(Note if leg cannot abduct and ext. rot. to contact surface in supine)

L □ R □ Shoulder protraction

L □ R □ Elbow flexion

L □ R □ Neck rotation

L □ R □ Neck lateral flexion

□ Plagiocephaly

□ Fixed spinal curve

Behavioral State : (Brazelton TB, Neonatal Behavioral Assessment Scale, 2nd ed.,1984)

State 1 Deep sleep

State 2 Light sleep

State 3 Drowsy or semi-dozing

State 4 Alert, with bright look

State 5 Eyes open, considerable activity

State 6 Crying

Testing environment:

Ideally test first thing in the AM or same time of day about 1 hour after feeding

Test on a firm padded mat

Diaper /onesie only unless the infant is cold

Test with toys

May use pacifier only if needed to maintain state 4 or 5 (see definition).

Mark as CNT (could not test) if patient could not be tested DO NOT MARK 0

_____ _____

SIGNATURE DATE

21.3. Gross Motor Skills Checklist

Gross Motor Skills Checklist

Subject: Date: Evaluator:

Item Description	Image	Able to Perform?	
1. Hands to midline in supine		Yes	No
2. Holds head upright while carried		Yes	No
3. Lifts head in prone		Yes	No
4. Rolls from side to back		Yes	No
5. Rolls from back to side		Yes	No
6. Props on hands in prone		Yes	No
7. Sits with support		Yes	No
8. Rolls from back to stomach		Yes	No
9. Rolls from stomach to back		Yes	No
10. Sits without support		Yes	No

Item Description	Image	Able to Perform?	
11. Sits without support and plays		Yes	No
12. Holds hands and knees position		Yes	No
13. Pulls to Stand		Yes	No
14. Cruises at support surface		Yes	No
15. Walks with support		Yes	No
16. Takes independent steps		Yes	No

21.4. Detailed Procedures for CMAP and MUNE

Participants were evaluated at baseline, 1 month post-infusion, 3 months post-infusion, and every three months thereafter for up to 24 months. All electrophysiological studies were performed by a single electromyographer (WDA) certified in Electrodiagnostic Medicine by the American Board of Electrodiagnostic Medicine. Prior to the start of the study, a standardized and technically detailed compound motor action potential (CMAP) and Motor Unit Number Estimation (MUNE) protocol was developed based upon best available evidence for incorporation into the study. A single digital EMG machine was used for all recordings. Electrophysiological measures included peroneal CMAP from the tibialis anterior and ulnar CMAP and MUNE from the abductor digiti minimi. During the recordings, temperature was measured at the hand and leg using an infrared temperature probe (ADD Company). If temperature was <33 degrees centigrade, a warming pack was used to warm the skin to ≥33 degrees centigrade prior to collecting data. Disposable ring electrodes (ADD electrode/company) were used for the G1 and G2 recording electrodes during the CMAP and MUNE responses. The length of the ring electrodes were trimmed with scissors to match the width of the muscle being tested.

For the peroneal CMAP, the G1 electrode was placed orthogonally at the proximal third of the tibialis anterior muscle. The G2 electrode was placed on the distal leg at the anterior aspect of the ankle. A ground electrode was placed between the point of stimulation and the G1 electrode. The peroneal nerve was supramaximally stimulated at or proximal to the fibular head. Maximum CMAP amplitude was recorded in mV with a lower limit of detection of 0.1 mV. Area was measured using the initial negative peak in mVsec. For the Ulnar CMAP and MUNE recordings, the G1 electrode was placed orthogonally over the abductor digiti minimi muscle at the proximal third of the distance between the fifth metacarpophalangeal joint and the pisiform bone. The G2 was placed on the medial aspect of the fifth metacarpophalangeal joint. The ground electrode was placed between the point of stimulation and the G1 electrode. The ulnar nerve was supramaximally stimulated at the ulnar groove. CMAP maximum amplitude and area were recorded with area being measured only for the initial negative peak. Supramaximal CMAP response was recorded using a MUNE protocol in the EMG system which allowed measurement in μV. Then, single motor unit potential (SMUP) responses were visualized using a live running screen and recorded. The amplitudes (in μV) of up to ten distinct, recurrent SMUP responses were recorded and stored. The MUNE values were calculated using the ulnar CMAP amplitude divided by average SMUP amplitude