Protocol Number: AVXS-101-CL-302

Official Title: Phase 3, Open-Label, Single-Arm, Single-Dose Gene Replacement Therapy Clinical Trial for Patients with Spinal Muscular Atrophy Type 1 with One or Two SMN2 Copies Delivering AVXS-101 by Intravenous Infusion

NCT Number: NCT03461289

Document Date: 03-Nov-2020

Protocol Number and Title:	**AVXS-101-CL-302** **Phase 3, Open-Label, Single-Arm, Single-Dose Gene Replacement Therapy Clinical Trial for Patients with Spinal Muscular Atrophy Type 1 with One or Two *SMN2* Copies Delivering AVXS-101 by Intravenous Infusion**
Protocol Version and Date:	**Version 7.0 Amend 6.0 / 25 Jun 2020**
Author(s):	[redacted] **Principal Biostatistician**
SAP Version:	**2.0**
SAP Version Date:	**03-Nov-2020**

Revision History

Version #	Date	Document Owner	Revision Summary
0.1	29-Jan-2018	[redacted]	Initial Release Version
0.2	12-Mar-2018	[redacted]	Second draft
1.0	10-May-2018	[redacted]	Final Version
2.0	03-Nov-2020	[redacted]	Update due to protocol amendment and COVID -19 impact

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

APPROVALS

Principal Biostatistician [redacted] AveXis, Inc	Date (dd-Mmm-yyyy)
Chief Medical Officer [redacted] AveXis, Inc.	Date (dd-Mmm-yyyy)
Chief Regulatory and Quality Officer [redacted] AveXis, Inc.	Date (dd-Mmm-yyyy)
Head of Biostatistics [redacted] AveXis, Inc.	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
CK-MB	Creatine kinase isoenzyme
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
ELISAs	Enzyme-Linked Immunosorbent Assays
ELISpot	Enzyme-Linked ImmunoSpot
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
MAR	Missing at random
Max	Maximum
MCAR	Missing completely at random
MedDRA	Medical Dictionary for Regulatory Activities
Min	Minimum
N/A	Not Applicable
NA	Not Applicable
NCI	National Cancer Institute
NTF	Note-To-File
OR	Observational Research
PASS	Post Authorization Safety Study
PAES	Post Authorization Efficacy Study
PBMC	Peripheral Blood Mononuclear Cells

Abbreviation	Description
PCS	Potentially Clinically Significant
PD	Pharmacodynamics
PDS	Pharmacodynamics Set
PK	Pharmacokinetic
PKS	Pharmacokinetic Set
PNCR	Pediatric Neuromuscular Clinical Research Network
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
SFC	Spot forming cells
SI	Standard International System of Units
SMN	Survival Motor Neuron
SMQ	Standardized MedDRA Queries
SOC	System Organ Class
SOP	Standard Operating Procedure
SS	Safety Set
TEAE	Treatment Emergent Adverse Event
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
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 + 1)/ 30. Age at Dosing = (Date of Gene Therapy Infusion – Date of Birth + 1)/ 30.
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.

Abbreviation	Description
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.
Day 1	The day of the gene therapy treatment.
Dose	Total vector genome administrated (in $^{vg}/_{\text{patient weight on Day } -1(\text{kg})}$).
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.
CTCAE	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.
PNCR	Pediatric Neuromuscular Clinical Research Network
Study Day	For any event of interest, Study Day = calendar date of event – calendar date of gene transfer +1. Day 1 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

The purpose of this document is to provide further details about the statistical analysis methods, data derivations and data summaries to be employed in the study protocol 15699-AVXS-101-CL-302: *Phase 3, Open-Label, Single-Arm, Single-Dose Gene Replacement Therapy Clinical Trial for Patients with Spinal Muscular Atrophy Type 1 with One or Two SMN2 Copies Delivering AVXS-101 by Intravenous Infusion*. This statistical analysis plan (SAP) has been based on International Conference on Harmonization (ICH) E3 and E9 guidelines and in reference to protocol version 2.1: dated 09 January 2018 and Annotated Case Report Form (aCRF): dated 2 November 2017. The statistical analysis plan covers statistical analysis, tabulations and listings of all data including effectiveness and safety data. Analyses will be performed using SAS® Version 9.3 (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.

2.1. Responsibilities

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

[redacted] is responsible for deriving the data set according to CDISC standards and creating data set specifications based on the SAP. [redacted] will perform the statistical analyses and are responsible for the production and quality control of all tables, listings and figures.

2.2. Timing of Analysis

The primary analysis will occur at such time that the last intent-to-treat (ITT) patient has reached 18 months of age; database lock will occur after all enrolled patients have completed the 18 months of age visit or have discontinued study early.

2.2.1. DSMB Quarterly Data Reviews

The DSMB's primary responsibilities will be to:

- safeguard the interests of study patients and assess the safety of the study treatment(s) and study procedures in a confidential manner
- assess the data quality, completeness and timeliness and provide recommendations about stopping or continuing the studies or otherwise modifying the studies
- contribute to enhancing the integrity of the studies, by formulating recommendations relating to the patient recruitment, selection and patient management during the studies
- further enhance the ability to evaluate the cumulative safety data by making recommendations regarding the format of the statistical summaries in the open and closed reports
- attend to safety and risk: benefit considerations in tandem with study designs, planned objectives, and patient care in accordance with ICH, FDA, EMA, Declaration of Helsinki, and Operational Guidelines for Establishment and Function of Data Safety Monitoring Boards/Data Monitoring Committees

- consider factors external to the studies when relevant information becomes available, such as scientific or therapeutic developments that may have an impact on the safety of the participants or the ethics of the studies
- review the conduct of the studies including protocol violations

The DSMB will serve in an advisory role to the AveXis Study Team. AveXis representatives will be responsible for promptly reviewing the DSMB recommendations, determining whether any changes to the studies are required (including protocol amendments), and initiating any changes to the studies.

2.2.2. DSMB Reporting and Meetings

Reports describing the status of the study will be prepared by AveXis and sent to each DSMB member prior to each meeting for review and preparation. DSMB meetings will occur on quarterly basis, aligning with the AveXis quarterly data cut reviews (as AVXS summary report completed in following month).

Reports include 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 grade 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) has been 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 Syneos Safety Management Plan (SMP). The DSMB could recommend early termination of the trial for reasons of safety. Study enrollment could be halted by the Sponsor when any patient experienced a Grade 3 or higher adverse event toxicity that is possibly, probably, or definitely related to the study drug. This includes any patient death, important clinical laboratory finding, or any severe local complication in the injected area related to administration of the study agent.

The DSMB charter and all materials developed for the DSMB to review and all documented recommendations will be maintained according to confidentiality requirements.

3. STUDY OBJECTIVES

3.1. Primary Objective(s)

The primary objective is to:

- Determine efficacy by demonstrating achievement of developmental milestone of sitting without support for at least 10 seconds up to 18 months of age as defined by WHO Motor Developmental Milestones

3.2. Secondary Objective(s)

The secondary objective is to:

Determine efficacy based on survival at 14 months of age. Survival is defined by the avoidance of combined endpoint of either (a) death or (b) permanent ventilation which is defined by tracheostomy or by the requirement of ≥ 16 hours of respiratory assistance per day (via non-invasive ventilatory support) for ≥ 14 consecutive days in the absence of an acute reversible illness, excluding perioperative ventilation. Permanent ventilation, so defined, is considered a surrogate for death.

3.3. Exploratory Objectives

The exploratory objectives, assessed at any visit up to and including the 18 month of age visit, are to:

- Determine the efficacy of AVXS-101 by demonstrating the achievement of developmental milestone of ability to hold head erect without support as defined by BSIDv03 Gross Motor Subtest Item #4
- Determine the efficacy of AVXS-101 by demonstrating the achievement of developmental milestone of ability to roll over as defined by BSIDv03 Gross Motor Subtest Item #20
- Determine the efficacy of AVXS-101 by demonstrating the achievement of developmental milestone of functional independent sitting for at least 30 seconds as defined by BSIDv03 item #26
- Determine efficacy by demonstrating achievement of developmental milestone of ability to stand with assistance for at least 2 seconds as defined by BSIDv03 Gross Motor Subtest Item #33.
- Determine efficacy by demonstrating achievement of developmental milestone of ability to crawl forwards on hands and knees for a distance of at least 5 feet as defined by BSIDv03 Gross Motor Subtest Item #34.
- Determine the efficacy of AVXS-101 by demonstrating the achievement of developmental milestone of ability to pull to stand as defined by BSIDv03 Gross Motor Subtest Item #35
- Determine efficacy by demonstrating achievement of developmental milestone of ability to walk with assistance as defined by BSIDv03 Gross Motor Subtest Item #37.

- Determine efficacy by demonstrating achievement of developmental milestone of ability to stand alone for at least 3 seconds as defined by BSIDv03 Gross Motor Subtest Item #40.
- Determine efficacy by demonstrating achievement of developmental milestone of ability to walk at least 3 steps without support as defined by BSIDv03 Gross Motor Subtest Item #42.
- Determine the efficacy of AVXS-101 by demonstrating the achievement of developmental milestone of ability to crawl as defined by WHO development milestones
- Determine the efficacy of AVXS-101 by demonstrating the achievement of developmental milestone of ability to stand with assistance as defined by WHO development milestones
- Determine the efficacy of AVXS-101 by demonstrating the achievement of developmental milestone of ability to stand alone as defined by WHO development milestones
- Determine the efficacy of AVXS-101 by demonstrating the achievement of developmental milestone of ability to walk with assistance as defined by WHO development milestones
- Determine the efficacy of AVXS-101 by demonstrating the achievement of developmental milestone of ability to walk alone as defined by WHO development milestones
- Determine the efficacy of AVXS-101 by demonstrating improvement of motor function as determined by improvement of the raw score from baseline to the maximum using fine and gross motor domain components of the Bayley Scales of Infant and Toddler Development (Version 3)
- Determine the efficacy of AVXS-101 by demonstrating improvement of gross motor function as determined by change from baseline to maximum in Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders (CHOP-INTEND) score
- Determine the effect of AVXS-101 on the ability to remain independent of ventilator support, defined as requiring no daily ventilator support/usage at 18 months of age.
- Determine effect on the ability to thrive defined as achieving all of the following:
  - Does not receive nutrition through mechanical support (i.e., feeding tube)
  - Ability to tolerate thin liquids as demonstrated through a formal swallowing test
  - Maintains weight (> third percentile for age and gender)
- Age at which patient first achieves independent sitting (10 seconds)

3.4. Safety Objectives

The safety objectives are to:

- Evaluate the safety of AVXS-101 in patients with SMA Type 1

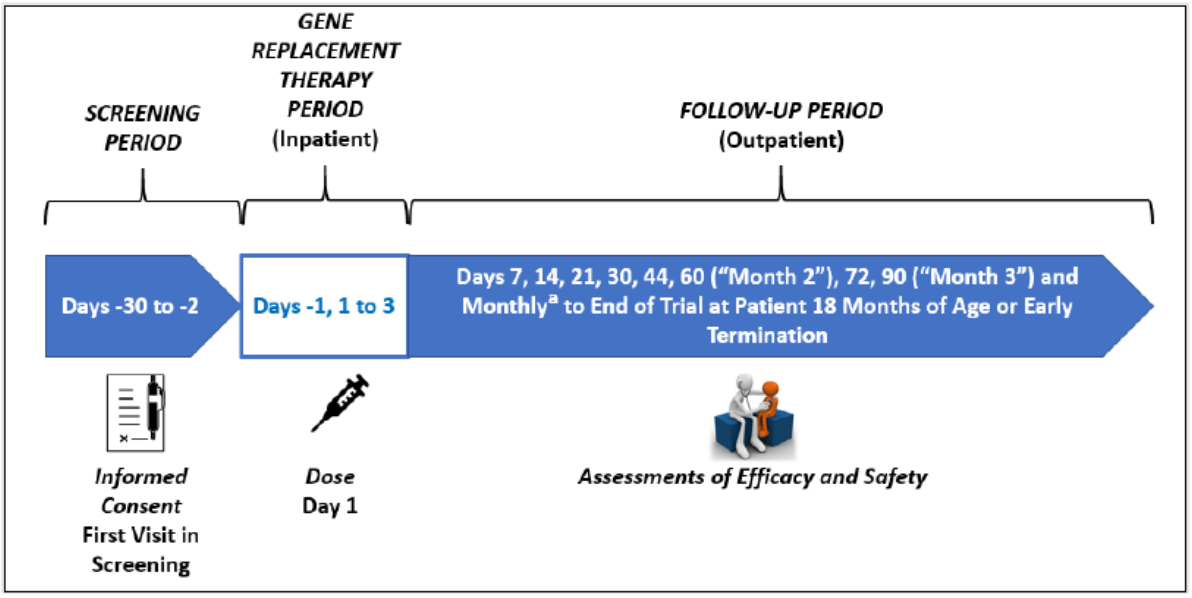
3.5. Study Design

This is a Phase 3, open-label, single-arm, single-dose trial of AVXS-101 (gene replacement therapy) in patients with SMA Type 1 with one or two copies of *SMN2*. Up to 30 patients < 6 months (< 180 days) of age **at the time** of gene replacement therapy (Day 1) will be enrolled.

The trial includes 3 trial periods: screening, gene replacement therapy, and follow-up (**Error! Reference source not found.**1). During the screening period (Days –30 to –2), patients whose parent(s)/legal guardian(s) provide informed consent will undergo screening procedures to determine eligibility for trial enrollment. Patients who meet the entry criteria will enter the in-patient gene replacement therapy period (Day –1 to Day 3). On Day –1, patients will be admitted to the hospital for pre-treatment baseline procedures. On Day 1, patients will receive a one-time intravenous (IV) infusion of the equivalent of AVXS-101 cohort 2 dose received in the AVXS-101-CL-101 trial over approximately 30-60 minutes, dependent upon required total volume, and will undergo in-patient safety monitoring over the next 48 hours. Patients may be discharged 48 hours after gene replacement therapy, based on Investigator judgment. During the outpatient follow-up period (Days 4 to End of Trial at 18 months of age), patients will return at regularly scheduled intervals for efficacy and safety assessments until the patient reaches 18 months of age. Any missed visit should be rescheduled as soon as possible, but within the visit windows specified in the schedule of assessments.

All post-treatment visits will be relative to the date on which gene replacement therapy is administered, until the patient is 14 months of age, after which will be relative to the patient's date of birth. For the 14 and 18 months of age visits, the patient will return within 0 to 14 days **after** the date on which the patient reaches 14 and 18 months of age, respectively. The 18 months of age visit will also serve as the End of Trial visit. After the End of Trial visit, eligible patients will be asked to participate in the long-term follow-up trial.

Figure 1: Study Design Schematic



Note: After the End of Trial visit at 18 months of age or at the time of early discontinuation patients will be invited to participate in a long term follow up study conducted under a separate protocol.

[a] All post-treatment visits will be relative to the date on which gene replacement therapy ("dose") is administered until the patient is 14 months of age, after which all visits will be relative to the patient's date of birth. *Note: Depending on the patient's age at dosing, the duration of participation at the end-of-trial visit can vary from approximately 12 months (baby dosed at approximately 6 months of age) to approximately 18 months (baby dosed near birth "0 months of age").*

A schedule of trial assessments is provided in Section 3.5.1.

3.5.1. Schedule of Assessments

Table 1: Schedule of Assessments

The blue highlighted section numbers are referring the sections in the protocol version 7.0, amendment 6.0 25 Jun 2020 only.

Trial Period	Screening	Gene Replacement Therapy (In-patient)				Follow-up (Outpatient)										Notes
Nominal Study Day (Visit Relative To Dose)	–30 to –2	-1	1	2	3	7	14	21	30	44	60 (or Month 2)	72	90 (or Month 3	Month 4,5,7,8,10,11,13	Month 6,9,12	End of Trial (18 months of age or ET)	Depending on the patient's age at dosing, the duration of participation can vary from approximately 12 months (baby dosed at approximately 6 months of age) to approximately 18 months (baby dosed near birth, "0 months of age").
Visit According To Months of Age														14,16,17	15		
Window	(No Window)					± 2 days					± 7 days (Except +0–14 days at 14 Months of Age)					+0–14 days	
Informed Consent	X																See Section 18.3.
AVXS-101 Infusion			X														See Section 10.6. Day 1 assessments will be performed prior to the start of gene replacement therapy infusion.
BSIDv03/WHO Developmental Milestones (with video)	X								X		X		X	X	X	X	See Sections 11.1, 11.2.1, 11.2.2 and 11.3.
BSIDv03 Gross and Fine Motor Subtests (with video)	X								X		X		X	X	X	X	See Sections 11.1, 11.2.1, 11.2.2 and 11.3.
CHOP INTEND (with video)	X	X				X	X	X	X		X		X	X	X	X	See Sections 11.2.3 and 11.3. Patients who achieve 3 consecutive CHOP INTEND scores ≥ 58 will not continue CHOP INTEND assessments.

Trial Period	Screening	Gene Replacement Therapy (In-patient)				Follow-up (Outpatient)										Notes
Nominal Study Day (Visit Relative To Dose)	–30 to –2	-1	1	2	3	7	14	21	30	44	60 (or Month 2)	72	90 (or Month 3	Month 4,5,7,8,10,11,13	Month 6,9,12	End of Trial (18 months of age or ET)	Depending on the patient's age at dosing, the duration of participation can vary from approximately 12 months (baby dosed at approximately 6 months of age) to approximately 18 months (baby dosed near birth, "0 months of age").
Visit According To Months of Age														14,16,17	15		
Window	(No Window)					± 2 days					± 7 days (Except +0–14 days at 14 Months of Age)					+0–14 days	
Demographic/Medical History	X																See Section 12.1.1.
Physical Exam	X		X	X	X	X	X	X	X		X		X	X	X	X	See Section 12.1.2.
Vital Signs/Weight & Length	X	X	(X)	X	X	X	X	X	X		X		X	X	X	X	See Section 12.1.3. (X): Vital signs will be continuously monitored throughout the infusion of gene replacement therapy and recorded every 15 minutes (+/- 5 minutes) post dose for the first 4 hours after the start of infusion, then every hour (+/- 15 minutes) until 24 hours after the start of infusion. Axillary temperature will be recorded pre- and post-infusion.
12-Lead ECG	X			X					X				(X)	(X)	(X)	X	See Sections 12.1.4, 12.1.5, and 12.1.6.
Echocardiogram	X								X				(X)	(X)	(X)	X	(X): Completed every 3 months, starting at Month 3 until 12 months post-dose.
24-hour Holter Monitor	X	X	X	X	X				X		X		(X)	(X)	(X)	X	
Pulmonary Examination	X	X		X	X	X	X	X	X		X		X	X	X	X	See Section 12.1.7.
Swallowing Test	X														(X)	(X)	See Section 12.1.8.

Trial Period	Screening	Gene Replacement Therapy (In-patient)				Follow-up (Outpatient)											Notes
Nominal Study Day (Visit Relative To Dose)	–30 to –2	-1	1	2	3	7	14	21	30	44	60 (or Month 2)	72	90 (or Month 3	Month 4,5,7,8,10,11,13	Month 6,9,12	End of Trial (18 months of age or ET)	Depending on the patient's age at dosing, the duration of participation can vary from approximately 12 months (baby dosed at approximately 6 months of age) to approximately 18 months (baby dosed near birth, "0 months of age").
Visit According To Months of Age														14,16,17	15		
Window	(No Window)					± 2 days					± 7 days (Except +0–14 days at 14 Months of Age)					+0–14 days	
																	(X): Completed every 6 months, starting at Month 6 through the End of Trial at 18 months of age.
Photograph of Infusion Site			(X)	X	X	X	X	X	X								See Section 12.1.9. (X): Day 1 infusion site photograph will be performed prior to the start of gene replacement therapy infusion.
Hematology/Chemistry /Urinalysis	X	(X)		X		X	X	X	X	(X)*	X	(X)*	X	X	X	X	See Sections 12.1.10.1, 12.1.10.2 and 12.1.10.3. (X): Laboratory samples collected on Day -1 to be processed locally, prior to dosing. (X)*: Liver function test (AST, ALT, total bilirubin, direct bilirubin, alkaline phosphatase, GGT) only.
CK-MB or Troponin I	X					X			X		X				X	X	See Section 12.1.10.2.
Virus Serology	X																See Section 12.1.10.4.
Capillary Blood Gas		X		X													See Section 12.1.10.5.
Immunological Testing: ELISA (anti-AAV9/SMN Ab)	X					X	X	X	X								See Section 12.1.10.6.

Trial Period	Screening	Gene Replacement Therapy (In-patient)				Follow-up (Outpatient)											Notes
Nominal Study Day (Visit Relative To Dose)	–30 to –2	-1	1	2	3	7	14	21	30	44	60 (or Month 2)	72	90 (or Month 3	Month 4,5,7,8,10,11,13	Month 6,9,12	End of Trial (18 months of age or ET)	Depending on the patient's age at dosing, the duration of participation can vary from approximately 12 months (baby dosed at approximately 6 months of age) to approximately 18 months (baby dosed near birth, "0 months of age").
Visit According To Months of Age														14,16,17	15		
Window	(No Window)					± 2 days					± 7 days (Except +0–14 days at 14 Months of Age)					+0–14 days	
AAV9 Ab Screen in Biological Mother	X																See Section 12.1.10.7.
Blood for Diagnostic Confirmation Testing	X																See Section 12.1.10.8.
Saliva, Urine, and Stool Samples (for viral shedding)	X			(X)	(X)	X	X	X	X								See Section 12.1.10.9. (X): Collected within 24 and 48 hours post dose.
Prophylactic Prednisolone		X	X	X	X	X	X	X	X	(X)	(X)						See Section 9.2.1. (X): After 30 days, prednisolone (or equivalent glucocorticoid) can be tapered for patients whose ALT values, and AST values are below the threshold.
Adverse Events	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	See Section 13.
Prior and Concomitant Medications	Collected from 2 weeks before gene replacement therapy until End of Trial visit																

Ab = antibody; ALT = alanine aminotransferase; AST = aspartate aminotransferase; BSIDv03 = Bayley Scales of Infant and Toddler Development version 3; CBC = complete blood count; CHOP INTEND = Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders; CK-MB = Creatinine kinase-muscle/brain; ECG = electrocardiogram; ELISA = enzyme-linked immunosorbent assay; ET = early termination; GGT = gamma glutamyl transferase; IRB = Institutional Review Board; LFT = liver function tests; WHO = World Health Organization

3.6. Patient Selection

The current version of the protocol (v7.0) allows for enrollment up to 30 patients, who are 6 months of age or younger at time of gene replacement therapy infusion (Day 1) diagnosed with SMA Type 1 who meet enrollment criteria, which may be either symptomatic or pre-symptomatic and are genetically defined by no functional survival motor neuron 1 gene (*SMN1)* as well as 1 or 2 copies of survival motor neuron 2 gene (*SMN2*). Enrolling up to thirty (30) patients under the broader enrollment criteria is projected to enable enrollment of at least twenty-five (25) patients that meet the Intent-to-Treat Population (ITT) criteria. The ITT population is identified as symptomatic patients with bi-allelic *SMN1* deletions and 2 copies of *SMN2* without the *SMN2* gene modifier mutation (c.859G>C) who meet all other study enrollment criteria. Patients with 1 copy of *SMN2*, pre-symptomatic patients and patients with the *SMN2* gene modifier mutation (c.859G>C) and other permutations outside of those specified in the ITT population will be evaluated separately as part of additional subgroup analyses. Details of all analyses will be contained within this Statistical Analysis Plan.

3.6.1. Inclusion Criteria

See protocol for details.

3.6.2. Exclusion Criteria

See protocol for details.

3.7. Determination of Sample Size

The current trial will enroll up to 30 patients. At least 25 symptomatic patients, the ITT population, will have biallelic deletion mutations of *SMN1,* two copies of *SMN2* and without the genetic modifier (c.859G>C), ensuring a reasonably robust safety database at the end of the study. The trial power is based upon efficacy analysis of the ITT population.

Based upon widely cited natural history of the disease (Pediatric Neuromuscular Clinical Research Network [PNCR]) [*Neurol*. 2014; 83(9):810-817], it is expected that no patients in this population would be expected to attain the ability to sit without support or accomplish other milestones (rolling over, standing, walking) prior to reaching 18 months of age. Assuming that the true response rate for the primary endpoint is actually zero (or as low as 0.1%) in the population of interest of the historical control, the primary efficacy endpoint hypothesis is:

$$\mathbf{H_0}: p_{\text{Avxs}-101} = 0.1\%$$

versus the alternative

$$\mathbf{H_a}: p_{\text{AVXS}-101} > 0.1\%,$$

where p is the proportion of functional independent sitting for at least 10 seconds up to and including the 18 month of age trial visit.

Based upon preliminary results from the ongoing Phase 1 clinical trial AVXS-101-CL-101, at least 50% of treated symptomatic patients with two copies of *SMN2* and without the genetic modifier (c.859G>C) are expected to achieve the primary efficacy endpoint of sitting without support for at least 10 seconds up to and including the 18 month of age trial visit. With the assumption for the true response rate of AVXS-101 being in the range of 20% - 50%, a sample size of 25 patients would provide power of greater than 95% to detect a significant difference with $\alpha = 0.025$ using a 1-sided exact test for a binomial proportion.

The secondary efficacy endpoint hypothesis:

$$\mathbf{H_0}: p_{AVXS-101} = p_{PNCR}$$

versus the alternative

$$\mathbf{H_a}: p_{AVXS-101} \neq p_{PNCR}$$

where p is the proportion of patients surviving at 14 months of age.

Based upon preliminary results from the same ongoing Phase 1 clinical trial AVXS-101, at least 80% of treated symptomatic patients with two copies of *SMN2* and without the genetic modifier (c.859G>C) are expected to achieve the secondary efficacy endpoint of survival at or beyond 14 months of age (based upon the definition of survival outlined above). It is anticipated that 75% of patients in the PNCR population would not survive beyond 13.6 months of age, and that these control patients will represent a 25% survival rate based upon the natural history of the disease. Given this efficacy, a sample size of 25 patients would provide power of >90% to detect a significant difference with α = 0.05 using a 2-sided Fisher's Exact test, as comparing to those control patients from existing natural history data sets (PNCR, NeuroNext [26]).).

3.8. Treatment Assignment and Blinding

This is an open-label study. All enrolled patients will receive AVXS-101.

3.9. Administration of Study Medication

Refer to Section 3.5 of the SAP.

3.10. Study Procedures and Flowchart

Refer to Section 3.5 of the SAP.

3.11. Statistical Hypotheses

3.11.1. Primary Efficacy Hypothesis

Primary and secondary efficacy analyses will be based on the ITT population, which consists of those symptomatic patients with bi-allelic deletions of *SMN1* and 2 copies of *SMN2* without the *SMN2* gene modifier mutation (c.859G>C). These analyses are to test the superiority of AVXS-101 versus the results from natural observation study (PNCR).

The primary efficacy hypothesis to be tested is:

$$\mathbf{H_0}: p_{AVXS-101} = 0.1\%$$

versus the alternative

$$\mathbf{H_a}: p_{AVXS-101} > 0.1\%,$$

where p is the proportion of functional independent sitting for at least 10 seconds up to an including the 18-month of age study visit.

3.11.2. Secondary Efficacy Hypothesis

The secondary efficacy hypothesis to be tested is:

$$\mathbf{H_0}: p_{AVXS-101} = pPNCR$$

versus the alternative

$$\mathbf{H_a}: p_{AVXS-101} \neq pPNCR$$

where p is the proportion 14 months of age event-free surviving.

Testing for the primary efficacy endpoint, independent sitting will first be performed using 1-sided exact binomial test. Only if the null hypothesis of equality in proportion of functional independent sitting is rejected at p <0.025, will the secondary efficacy endpoint survival improvement be tested using a 2-sided Fisher's exact test on the ITT population, comparing to patients from the natural observational study (PNCR) at 0.05 significance level. This hierarchy approach strongly prevents the Type I error rate from inflation.

4. ENDPOINTS

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

4.1.1. Primary Efficacy Endpoints

The primary efficacy endpoint is achievement of sitting without support for at least 10 seconds. This milestone will be assessed/determined by the qualified Clinical Evaluators at the investigational sites. The assessments will be captured on video from two camera angles during the protocol specified visits. Furthermore, videos recorded at home can also be considered in evaluating the developmental milestones efficacy endpoints. All videos, either captured at study visits or at home will then be reviewed and verified by an independent, external reviewer for concordance. Only milestones confirmed by the independent reviewer will be included for the primary analysis. Milestones determined by the qualified Clinical Evaluators will be presented in listings.

4.1.2. Secondary Efficacy Endpoints

The secondary efficacy endpoint is survival, defined as avoidance of either (a) death or (b) permanent ventilation through the 14 month of age visit. Permanent ventilation is defined by tracheostomy or requirement of ≥16 hours of respiratory assistance per day (via non-invasive ventilatory support) for ≥14 consecutive days in the absence of an acute reversible illness, excluding perioperative ventilation. Permanent ventilation, so defined, is considered a surrogate for death.

The respiratory assistance includes but is not limited to BiPAP machines, for example, Philips Trilogy 100 machine. Particularly, Trilogy 100 usage data will be saved in the SD card when patient uses the machine. Data in the SD card will be sent to AveXis by site during clinical visit. Daily usage information will be available when using DirectView software to extract the "compliance summary" from the raw data (.EDF files).

In addition, each site will be asked to verify whether permanent ventilation criteria were met for each patient. If a patient is reported to necessitate permanent ventilation, the date at which permanent ventilation was met will be reported by site.

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

For the following exploratory efficacy endpoints from achieving the ability to hold head erect to the ability to walk alone, only post-dosing achieved developmental milestones are considered. Milestones observed at baseline will not be counted in the summary table.

1. Proportion of patients that achieve the ability to hold head erect without support as defined by BSIDv03 Gross Motor Subtest Item #4
2. Proportion of patients that achieve the ability to roll over as defined by BSIDv03 Gross Motor Subtest Item #20
3. Proportion of patients that achieve functional independent sitting for at least 30 seconds as defined by BSIDv03 Gross Motor Subtest Item #26
4. Proportion of patients that achieve ability to stand with assistance for at least 2 seconds as defined by BSIDv03 Gross Motor Subtest Item #33

5. Proportion of patients that achieve to crawl forwards on hands and knees for a distance of at least 5 feet as defined by BSIDv03 Gross Motor Subtest Item #34
6. Proportion of patients that achieve to walk with assistance as defined by BSIDv03 Gross Motor Subtest Item #37
7. Proportion of patients that achieve to stand alone for at least 3 seconds as defined by BSIDv03 Gross Motor Subtest Item #40
8. Proportion of patients that achieve to walk at least 3 steps without support as defined by BSIDv03 Gross Motor Subtest Item #42
9. Proportion of patients that achieve the ability to crawl as defined by WHO Motor Developmental Milestones
10. Proportion of patients that achieve the ability to stand with assistance as defined by WHO development milestones
11. Proportion of patients that achieve the ability to stand alone as defined by WHO development milestones
12. Proportion of patients that achieve the ability to walk with assistance as defined by WHO development milestones
13. Proportion of patients that achieve the ability to walk alone as defined by WHO development milestones

- Improvement of raw score from baseline to the maximum using the Bayley Scales of Infant and Toddler Development (Version 3)

14. , Fine and Gross Motor Function Subtests
15. Maximum change from baseline to maximum CHOP-INTEND score
    - Proportion of patients achieving CHOP-INTEND score ≥40
    - Proportion of patients achieving CHOP-INTEND score ≥50
    - Proportion of patients achieving CHOP-INTEND score ≥58
16. Proportion of patients with the ability to remain independent of ventilator support, defined as requiring no daily ventilator support/usage at 18 months of age.
17. Proportion of patients maintaining the ability to thrive.
18. Age at which patient first achieves independent sitting (10 seconds).

4.2.1. CHOP-INTEND Score

There are 16 items in CHOP-INTEND assessment (Section **Error! Reference source not found.**). A rating of Brazelton behavioral states (Section **Error! Reference source not found.**) was recorded for each item. As suggested by manual [9], the optimal state for testing is state 4 (“alert, with bright look”) and 5 (“eyes open”). If a subject cannot be tested for an item due to an adverse behavioral state, it should be scored as “CNT” (cannot test) and NOT a zero. In item 1 to 11, 13 and 16, both left and right sides need to be evaluated, and the maximum score should be selected for the best score of the item. If both sides are scored “CNT”, the item should be scored as “CNT”. If only one side scored “CNT”, the other side score should be used for the best score

of the item. If any of these 16 items scored “Cannot test” or “CNT”, the total score should be set missing.

4.3. Safety Endpoints

Safety endpoint include physical examinations, pulmonary examinations, vital signs, capillary blood gas assessments, weight and length measurements, 12-lead ECGs, echocardiograms, 24-hour Holter monitors, swallowing tests, laboratory assessments, adverse event monitoring, and photographs of the infusion site. In general, safety assessments will be performed at the times scheduled in Table 1. All post-treatment visits are relative to the date on which gene replacement therapy is administered until the patient is 14 months of age, after which all visits will be relative to the patient’s date of birth.

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

4.4. Pharmacokinetic Endpoints

Not applicable for this SAP.

4.5. Pharmacodynamics 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

Length, weight, blood pressure, respiratory rate, pulse, body temperature, and pulse oximetry will be collected at every visit. For detailed schedule of events table, see Table 1.

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, dermatologic, lymphatic, and genitourinary.

4.6.4. Laboratory Evaluations

Blood will be collected throughout the study for standard blood chemistry and hematology tests as well as cardiac enzymes. Urine will be collected throughout the study for standard urinalysis exams.

4.6.5. Use of Non-Oral Feeding Support

A swallowing test will be performed at baseline and every 6 months to determine if the patient has signs of aspiration. If the test is positive for aspiration, the patient will be recommended 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.

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

A 12-lead electrocardiograms will be conducted on the schedule of assessment (Table 1). The ECG tracings or ECG machine data will be collected for centralized review and interpretation by a cardiologist.

A Holter monitor and serial ECG data collected at the scheduled time points. Twenty-four hour Holter monitoring will also be performed starting from Month 3 and every 3 months thereafter until 12 months post-dose and again at 18 months of age.

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 WHO Drug dictionary, most recent version available.

4.6.10. Non-invasive Ventilatory Support

Patients will be assessed by a pulmonologist at the time points specified in the schedule of assessments and may be fitted with a non-invasive positive pressure ventilator (e.g., Bilevel Positive Airway Pressure (BiPAP)) at the discretion of the pulmonologist and/or investigator. Non-invasive ventilatory support equipment may be provided by AveXis, Inc. through a third-party vendor (as necessary).

Each patient will be assessed by investigator whether permanent ventilation criteria has been met. If the patient meets the permanent ventilation criteria, the date of permanent ventilation met will be reported through [redacted]. Also, if [redacted] machine is used with SD card, the daily real hours of usage data will be sent to AveXis. Health-economics Endpoints

The analysis of healthcare resource utilization data may be performed for this study.

4.7. Other Endpoints

Additional analyses of the primary, secondary, and other additional endpoints utilizing a natural history cohort as a comparison may be incorporated, as appropriate.

5. ANALYSIS POPULATION

Safety will be assessed based on all patients who underwent gene therapy at day 1. The primary population for the primary and secondary efficacy analyses will be the ITT population. If enrolled, patients with 1 copy of *SMN2*, pre-symptomatic patients and patients with the *SMN2* gene modifier mutation (c.859G>C) and other permutations outside of those specified in the ITT population will be evaluated separately as part of additional subgroup analyses.

5.1. Safety Population

All patients who underwent gene therapy infusion will be included. All safety analyses will be based on the safety population, unless specified otherwise.

5.2. Intent-to-Treat (ITT) Population

The ITT population will consist of symptomatic patients with bi-allelic deletion of *SMN1* (exon 7/8 common homozygous deletions) and 2 copies of *SMN2* without the known gene modifier mutation (c.859G>C) who receive an IV infusion of AVXS-101 at less than 180 days of age.

5.3. Ability to Thrive ITT Population

The ability to thrive ITT population will consist of symptomatic patients with biallelic deletion mutations of *SMN1,* 2 copies of *SMN2* without the genetic modifier (c.859G>C), intact swallowing and receiving no enteral (mechanical) nutrition at baseline, who receive an IV infusion of AVXS-101 and have at least one post-baseline efficacy evaluation.

5.4. Efficacy Completers Population

The efficacy completers analysis population will consist of:

- All treated patients who reach 14 months of age for the survival endpoint or 18 months of age for the endpoint of achievement of sitting without support, OR
- All treated patients who meet discontinuation criteria, discontinue the study due to an AE, or experience death

5.5. All Enrolled Population

The all enrolled population will consist of all patients who receive an IV infusion of AVXS-101. Analyses of endpoints in this population are considered descriptive. For AVXS-101-CL-302, all enrolled population is the same as safety population.

5.6. PNCR Control Population

The PNCR Natural History dataset is drawn from a large natural history study (Finkel, 2014) performed on a large cohort of 337 patients with any form of spinal muscular atrophy followed at 3 large, internationally recognized tertiary medical centers with significant expertise in the management of SMA (Harvard University/Boston Children's Hospital, Columbia University and the University of Pennsylvania/Children's Hospital of Philadelphia). Previously identified patients followed in PNCR site clinics and newly diagnosed patients were enrolled. All eligible patients were offered participation in the PNCR study. Study visits were scheduled at baseline and at 2, 4, 6, 9, and 12 months and every 6 months thereafter. The SMA standard of care

guidelines (Wang, 2007) were used as a basis for providing uniform care among the study sites. For purposes of this study, sitting (for SMA Type 2) was defined as being able to sit independently for >10 seconds; children who were unable at any point to achieve this milestone were classified as SMA Type 1.

A natural history control population is drawn from the PNCR Natural History Dataset (Finkel, 2014) consisting of all patients with age of onset ≤6 months, bi-allelic deletion of *SMN1* (exon 7/8 common homozygous deletion) and 2 copies of *SMN2* for whom enrollment data (retrospective and prospective) is available to determine survival based on the defined criteria. The *SMN2* modifier mutation (c.859G>C) described by Prior and colleagues (Prior, Thomas W., et. al. "A positive modifier of spinal muscular atrophy in the *SMN2* gene." *The American Journal of Human Genetics* 85.3 (2009): 408-413.) was not assessed in the PNCR study cohort. Based on these criteria, 23 patients from the PNCR dataset are included in the control population. Among them, 16 patients reached the combined endpoint of death or the need for a minimum of 16 hours/day of noninvasive ventilation support for a minimum of 14 continuous days by 14 month of age and one patient discontinued the study at age of month 4. The survival rate at 14 months of age is approximately 25% (6/23), of which the control cohort will be used as comparator with ITT population for survival analysis.

5.7. Pharmacokinetic Set

Not applicable for this SAP.

5.8. Pharmacodynamics Set

Not applicable for this SAP.

5.9. Protocol Deviations

All deviations will be recorded in the [redacted] database and will be categorized in accordance with ICH E3 Guidelines. 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 one-sided with significance level α=0.025 or two-sided, at significance level α=0.05, 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 assumptions of parametric tests are not met.

Summary tables, listings, and figures, and statistical analyses will be done using SAS version 9.3 or higher. The primary and secondary efficacy analyses will be conducted on the ITT population. Safety analyses will be conducted on the Safety population.

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 1, 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.

The baseline values will be the first non-missing values collected for the control patients from the PNCR dataset.

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 1 is the day of the infusion of the gene therapy study drug.

6.2.3. Definition of Final Treatment Value

The first non-missing value on or after the participant reaches 14 months of age will be considered for the purposes of the secondary endpoint survival without permanent ventilation. There is no upper bound for the “14 months” age visit; that is, no visit is too late to represent the “14 months of age” analysis. The earliest visit on or after 420 days of age will be used in the “14 months of age” analysis.

Due to the COVID-19 pandemic, patients with the end of study (EOS) visit planned between Feb, 2020 and Sept, 2020 may not be able to complete the EOS visit in the protocol defined visit window. Based on current guidance, these patients will be allowed to complete the EOS when hospitals reopen for normal operation.

For developmental milestones, CHOP-INTEND score and Bayley scale, the final treatment value for each patient is defined as the first non-missing value on or after the participant reaches 18 months of age (540 days) with an upper limit of 570 days. Thus, the earliest visit between 540 and 570 days of age, inclusive, will be used in the "18 months of age" analysis. For other non-safety measures, the definition of final treatment value will be footnoted in the corresponding tables, listings or figures.

For safety data collected through the EOS visit will be used for safety analysis, no upper bound for final treatment value.

6.2.4. Derived and Transformed Data

6.2.4.1. Primary Efficacy Variables

The primary efficacy endpoint is of symptomatic SMA Type 1 patients with biallelic SMN1 deletions and 2 copies of SMN2 without the genetic modifier (c.859G>C) that achieve the ability to sit without support for at least 10 seconds at any visit up to and including the 18 months of age visit. It is defined by the World Health Organization Multicentre Growth Reference Trial (WHO MGRS), confirmed by video recording, as a patient who sits up straight with head erect for at least 10 seconds; child does not use arms or hands to balance body or support position.

6.2.4.2. Supportive (Secondary and exploratory) Efficacy Variables

The secondary efficacy endpoint, event-free survival, is defined as avoidance of any one of the following events up to 14 months of age, whichever occurred first:

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

Usage data for patients using non-invasive ventilatory support will be extracted directly from the device (e.g., Trilogy 100, Trilogy 200, ResMed Stellar) on an SD card, or USB stick and transferred as an external dataset, if available. Average daily use of non-invasive ventilatory support at baseline and monthly be summarized from data recorded on an SD card or USB stick and submitted in a format that is evaluable. If a patient did not have a device with the capability to record usage data at baseline and was using non-invasive ventilatory support prior to screening, the daily usage will be documented in the source documentation. 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 patients in each ventilation support category will be presented.

Supportive (exploratory) efficacy variables are as follows:

- CHOP-INTEND summary score: This is assessed by evaluating the Baseline CHOP-INTEND scores compared to each scheduled visit time points (see Table 4), collected from patient'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 determined for categorizing CHOP-INTEND summary scores: ≥40, ≥50, ≥58.
- Bayley Scales of Infant and Toddler Development© version 3: Two domains of motor and adaptive behavior will be assessed from patient's CRF at scheduled visits (see Table 1) with both raw and scaled scores provided.
- Age at which independent sitting (for 10 seconds) is first achieved.
- Ability to remain independent of ventilator support, defined as requiring no daily ventilator support/usage at 18 months of age.
- Ability to thrive as defined by maintaining the ability to swallow and to maintain weight (> third percentile based on WHO CGRS without need of gastrostomy or other mechanical or non-oral nutritional support at 18 months of age

6.2.4.3. First observed date of video confirmed milestone

All videos, either from study visits or caregiver provided home videos which meet WHO and BSID criteria for items on the developmental milestone checklist will be provided to an independent central reviewer for unbiased assessment of developmental milestone achievement. The independent central reviewer will document whether the video displays evidence of having achieved developmental milestone.

The date in which a Developmental Milestone is first observed and recorded on video, clinic or home video is entered into [redacted]. In the event that the Central Reviewer does not confirm the milestone, video evidence will be provided from subsequent clinic visit or home video footage until confirmation is received. The earliest date of video recording which is confirmed by central review is considered the date of developmental milestone achievement.

6.2.4.4. Safety Variables

The safety endpoint is defined as the patient meeting the following criteria:

- Any AE
- Any serious AE
- Any AE related to study product
- Any AE with CTCAE Grade ≥ Grade 3

This will be collected at all scheduled visits (see Table 1) from patient's CRF. Adverse Event collection methods for the observational natural history cohort were not consistent with rigorous Adverse Event collection methods demonstrated in the context of an interventional clinical research study; therefore, only descriptive statistics will be presented.

6.3. Handling Missing Data

All patients in the ITT analysis set will be included in the primary efficacy analysis for a given endpoint provided there are non-missing baseline values. If a patient misses a regularly scheduled visit, every attempt will be made to reschedule a visit within that visit window to obtain evaluations or tests.

For the primary endpoint (the ability to sit without support for 10 seconds at up to 18 months of age) collected during the study, only when a patient demonstrates the ability in a video recorded at a visit or at home and confirmed by a central reviewer, a "Yes/True" value will be reported in the MMI report. A missing value of the primary endpoint of a patient means the patient does not achieve the ability of the milestone. No imputation is necessary.

For secondary efficacy endpoint, only when a patient met the survival criteria, the date of death or permanent ventilation will be reported through [redacted] Patients who terminate the study prior to reaching 14 months of age for any reason will be considered treatment failures (event). Patients who terminate the study after 14 months of age but are still event-free surviving at the time of termination will be censored.

The following strategies will be implemented for different efficacy endpoints

- Central reviewed developmental milestone. No imputation will be implemented
- CHOP-INTEND score. No sub item score will be imputed. If the missing total score is due to "Can not test (CNT)", no imputation will be implemented. Otherwise, last observation carried forward (LOCF) will be applied to impute the missing values.
- Bayley Scales of Infant and Toddler Development. LOCF will be implemented to impute the missing values
- Swallowing, feeding support and ventilation support. No imputation will be implemented.

In addition, missing values for safety endpoints, including but not limited to AE, lab values, will not be imputed. For local lab, missing units or normal ranges will be reviewed by the medical team. If those missing information needs be imputed for selected lab tests, the imputed information will be captured in the Note to File (NTF). The following tables provides the most conservative way to impute the date of event, if a missing date observed and not able to be fixed on the source.

Table 2: Rules of Date Imputation: Pre-Dose Data

Partial Missing Start or Stop Date	Imputed Start Date	Imputed Stop Date
Missing month and day, but the year is present	January 1 of that year or dose date if the year is the same as the year of dose date	December 31 of that year
Missing day, but year and month are present	First day of that month or dose date if the year and month are the same as the year and month of dose date	Last day of that month
Missing month, but year and day are present	Missing month imputed as January	Missing month imputed as December

Table 3: Rules of Date Imputation: Post-Dose Data

Partial Missing Start or Stop Date	Imputed Start Date	Imputed Stop Date
Missing month and day, but the year is present	Date of dose	December 31 of that year
Missing day, but year and month are present	First day of that month or dose date if the year and month are the same as the year and month of dose date	Last day of that month
Missing month, but year and day are present	Date of dose	Missing month imputed as December

6.4. Visit Windows

Due to the visit schedule defined in the protocol, two different schedules are applied when creating statistical outputs including tables, figures and listings. One is based on study day post dosing and the other one is based on patient age.

6.4.1. Study Day Based Visit Windows

For efficacy analysis, the time windows specified in Table 4 describe how efficacy data will be assigned to protocol-specified time points, as displayed in the Schedule of Assessments in SAP Section 3.5.1. All time points and corresponding time windows are defined based on Dosing Date.

If more than one efficacy observation for a specific assessment is included in a time window, the assessment closer to the nominal time will be used. If there are two efficacy observations equally distant to the nominal time, the latest one will be used in analyses. Any efficacy assessments occurring outside the analysis windows will be considered an assessment of an unscheduled visit.

Table 4: Study Day Based Visit Window

Scheduled Visit	Nominal Days (Study Day)	Acceptable Analysis Window in Study Days (Min Day – Max Day)
Baseline	-	<1
Month 1	30	1 to 44
Month 2	60	45 to 74
Month 3	90	75 to 104
Month 4	120	105 to 134
Month 5	150	135 to 164
Month 6	180	165 to 194
Month 7	210	195 to 224
Month 8	240	225 to 254
Month 9	270	255 to 284
Month 10	300	285 to 314
Month 11	330	315 to 344
Month 12	360	345 to 374
Month 13	390	375 to 404
Month 14	420	405 to 434
Month 15	450	435 to 464
Month 16	480	465 to 494
Month 17	510	495 to 539
Month 18	540	540 to 569
Month K	K*30	K*30 to (K+1)*30-1
Final Treatment Visit and 14 Months of Age Analysis Window		0 to ≤14 days after patient reaches 14 months of age (420-434 days of age) visit window; 0 to ≤14 days after patient reaches 18 months of age (540 days of age) visit window

*Depending on age at time of dosing, patients will complete varying number of visits. The patient must complete a visit within 2 weeks after reaching 14 months of age; subsequent visits will be scheduled from that visit date. The patient must also complete the final study visit within 2 weeks after reaching 18 months of age.

Due to COVID-19 pandemic, a few patients may complete the final visit (18 months of age) months outside of protocol-defined visit window. Summary tables for efficacy analysis will be created based on the visits up to month 18 of age and any data collected after month 18 of age visit to the end of study visit will be grouped into “Month 19 of age and older. Data collected through the final visit will be presented in the listings. In case, multiple record falls within “Month 19 of age and older” visit then first record will be used for analysis.

Safety data, such as laboratory results, vital signs, ECGs, and physical exams will be assessed by date and study day. For lab data collected on an unscheduled visit, the protocol defined visit windows will be applied. For change from baseline analyses the value associated with the scheduled visit will be used. Due to the visit window (± 7 days) for Day 60 visit and Day 72

visit, there is an overlap when assigning data to these two visits. Data collected between Day 53 to Day 66 will be assigned to Day 60 visit. Data collected between Day 67 to Day 79 will be assigned to Day 72 visit. For summaries of shifts from baseline and potentially significant values, all post baseline values will be considered for these analyses. Baseline for safety measures will be defined as the latest value before Day 1.

6.4.2. Age Based Visit Window

Summary of efficacy endpoints, including CHOP-INTEND score and Bayley Scales will also be presented by Age (in months) which assigned as in Table 5. If more than one efficacy observation for a specific assessment is included in a time window, the assessment closer to the nominal time will be used. If there are two efficacy observations equally distant to the nominal time, the latest one will be used in analyses.

Table 5: Age Based Window in Months

Age (Months)	Window for Days since birth
< 1 month	2 to 29
1 month	30 to 59
2 months	60 to 89
3 months	90 to 119
...	...
15 months	450 to 479
16 months	480 to 509
17 months	510 to 539
18 months	540 to 569
K months	K*30 to (K+1)*30-1

6.5. Data Cutoff Rule and Date

Due to COVID pandemic, a few patients delayed the scheduled end of study visits (month 18 of age visit). Analyses will be done after database lock.

6.6. Pooling of Centers

This is a multiple-center single dose treatment study. Given that the expected enrollment at most of sites is expected to be fewer than 5 patients, patients from all sites will be pooled together into a combined site in the analyses for the primary and secondary effectiveness endpoints. A site effect will be not examined because there will be too few patients within each center to provide an informative estimate of this effect.

6.7. Subgroups

Two subgroups based on age at dosing (study day 1) will be assessed in the selected analysis (developmental milestones and statistical modeling parts):

- Patients with age at dosing (study day 1) <= 4 months (120 days)

- Patients with age at dosing (study day 1) > 4 months (120 days).

7. DEMOGRAPHIC, OTHER BASELINE CHARACTERISTICS AND MEDICATION

7.1. Patient Disposition and Withdrawals

Patient disposition will be presented for all patients, which will include the following:

- The number of patients screened
- The number (%) of screened patients who screen failed and the reasons for screen failure (inclusion/exclusion criteria, withdrew consent, and/or other) will be summarized. A listing of reason for screen failure will be provided for all patients who screen failed.
- The number (%) of patients in the safety population
- The number (%) of patients in the Efficacy Completer Population
- The number (%) of patients reach the 14 months of age for survival endpoint
- The number (%) of patients who completed the study, defined as all patients with a final scheduled visit conducted either remotely or at the study site
- The number (%) of patients who discontinued from the study and the associated reasons

For patients who discontinued due to AE, a separate summary table will be produced in order to detail the type of AE that resulted in the discontinuation (whether serious fatal, non-serious fatal, or non-fatal).

7.2. Demographic Data

The age of the patient at the time of gene therapy infusion will be summarized. The distribution of patients by sex will be presented. Patient demographics will be summarized using the Safety population.

Demographic data will be determined using the following calculations:

Age at Study day 1 = (Study day 1 visit date - date of birth + 1), expressed in days

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

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

The following statistics regarding the patient'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. In particular, the following parameters will be summarized regarding symptoms and history of Spinal Muscular Atrophy: age at symptom onset (if applicable), baseline SMA symptoms (if applicable), family history of SMA, and number of siblings affected by SMA.

Patient baseline characteristics will be summarized on the Safety population by subgroup and overall.

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 patients in the safety analysis set with a particular condition/diagnosis will be summarized overall. Patients 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 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 patients in the safety analysis set taking prior or concomitant medications will be summarized by generic drug name based on the WHO Drug Dictionary.

7.4.1. Specific Medication Subgroups

To reduce the host immune response to the AAV-based therapy, patients will receive prophylactic prednisolone (approximately 1 mg/kg/day) 24 hours prior to the gene transfer and continuing for approximately 30 days. After 30 days of treatment, the dose of prednisolone can be tapered for patients whose ALT values, AST values, and T-cell response are ≤2 X ULN for ALT and AST, and <100 SFC/10^6 PBMCs in accordance with the following treatment guideline:

- Weeks 5 and 6: 0.5 mg/kg/day
- Weeks 7 and 8: 0.25 mg/kg/day
- Week 9: prednisolone discontinued

The total number of days receiving prophylactic prednisolone and total cumulative dose of prophylactic prednisolone administered during the study (mg/kg, derived from the prednisolone log) will be computed for each patient. In case when other glucocorticosteroids were used, the doses should be converted into Prednisolone doses according to the conversion ratio, for example 4:1 for Hydrocortisone.

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 patient 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 patient,

$$\begin{aligned}\text{Total Cumulative Dose} = &(1.0 \text{ mg/kg} \times 19 \text{ days}) + \\ &(2.0 \text{ mg/kg} \times 10 \text{ days}) + \\ &(1.5 \text{ mg/kg} \times 5 \text{ days}) + \\ &(1.0 \text{ mg/kg} \times 5 \text{ days}) + \\ &(0.5 \text{ mg/kg} \times 6 \text{ days}) \\ = &\ 54.5 \text{ mg/kg}\end{aligned}$$

Exposure will be summarized for the Safety Analysis Set.

7.4.2. Other Therapies

Non-medication Therapies/Procedures will be defined as “Prior” and/or “Concomitant”. Prior Non-medication Therapies/Procedures are defined as therapies or procedures started prior to injection of AVXS-101. Concomitant Non-medication Therapies/Procedures are defined as therapies or procedures ongoing at time of injection of AVXS-101 or started after the injection. Non-medication Therapies/Procedures will not be summarized. A listing will be provided only.

8. EFFICACY

8.1. Primary Efficacy Endpoints and Analysis

8.1.1. Primary Efficacy Endpoint: Sitting Without Support up to 18 months of age

An independent expert reviewer will evaluate recorded videos of WHO MGRS assessments conducted at clinic visits. The reviewer will judge whether the video reveals evidence of the milestone achievement of functional independent sitting (Child sits alone without support for at least 10 seconds).

During the Screening visit, the physical therapist will complete an assessment of baseline milestone achievement, including functional independent sitting; this assessment must address all milestones/items noted in Appendix 17.1 that are at or below the child's expected function for age. Items that are below the child's expected function for age that are not successfully achieved during the baseline evaluation should be repeated at subsequent visits until successfully performed.

The milestone of sitting without support will be assessed at every visit regardless of their initial status.

The number and percent of patients whom, through video evidence, exhibit the milestone achievement of sitting without support at any visit up to and including18 months of age study visit will be summarized for the ITT population (and for identified subgroups). A one-sided Exact Binomial Test will be used to test the null hypothesis of p=0.1% at significance level of 0.025. Furthermore, the corresponding 97.5% confidence intervals will be estimated by the exact method for binomial proportions.

Descriptive statistics and statistical analyses for subgroups specified in Section 6.7 will be provided to illustrate the consistency of the response rate across subgroups. If too few patients are in the subgroup to do an analysis, or an appropriate natural history or other appropriate comparator is not available, and a change from baseline analysis is not appropriate, descriptive statistics only will be provided.

This analysis will be repeated for the efficacy completers population.

An additional set of sensitivity analyses will be conducted to include patients who demonstrated the milestone of sitting without support, which occurred after but outside of the SAP-defined visit windows of the Month 18 of age visit due to COVID-19 disruptions.

8.2. Secondary Efficacy Endpoints and Analysis

8.2.1. Secondary Efficacy Endpoint: Avoidance of Death or Surrogate for Death (Permanent Ventilation)

The secondary efficacy endpoint is defined by avoidance of the combined endpoint of either (a) death or (b) permanent ventilation, defined as requirement of tracheostomy or ≥16 hours of 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.

Non-invasive ventilatory support usage data (only available from Trilogy 100) will be extracted directly from each patient's device; the data will be downloaded from the device's SD card through software provided by the manufacturer and transferred as an external dataset.

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 1, patient 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: The proportion surviving event-free to 14 months of age will be computed in the ITT population. Patients who terminate the study prior to reaching 14 months of age for any reason will be considered treatment failures (event).

As a comparator, in a natural history study of SMA Type 1 patients, Finkel et al (2014), 16 patients reached the combined endpoint of death or the need for a minimum of 16 hours/day of noninvasive ventilation support for a minimum of 14 continuous days by 14 month of age, one patient discontinued the study at age of month 4 and 6 patients event-free survived at age of 14 months.

The observed proportion surviving in the current study will be compared to the natural history data (5) of the matching cohort, using a 2-sided Fisher's Exact test, along with the corresponding 95% confidence intervals.

The number of patients survived in each subgroups specified in Section 6.7 will be provided.

Time to death or permanent ventilation through 14 months of age is an additional sensitivity analysis. In case there is at least 1 patient who is event-free, but discontinue study prior to 14-months of age, a sensitivity analysis, based on the approach for time to event, will be carried out. In this analysis, time to event will be censored at 14 months of age or at the date of

discontinuation. The survival rate at 14 months of age will be estimated by Kaplan-Meier method, and survival curves for AVXS treated patients and the PNCR cohort will be compared using log-rank test at significance level of 0.05.

Additionally, the proportion of patients who experience each of the following events by 14 months of age will be summarized:

- Death
- Permanent ventilation

8.3. Exploratory Efficacy Endpoint(s) and Analyses

The exploratory efficacy analyses will be based on Efficacy Completers Population. Due to COVID-19 impact, any efficacy data collected outside of the SAP-defined visit window of Month 18 of age visit (older than 570 days of age) will be grouped into "Month 19 of age or older" and included for summary tables unless specified. The hypothesis testing will be based on the data collected up to 18 months of age. Those data will also be included in the corresponding listing.

8.3.1. Proportion of Patients that Achieve Ability to Hold Head Erect Without Support

The ability to hold head erect without support is defined by BSIDv03 Gross Motor Subtest Item #4 and will be based upon the maximum function achieved at any time point up to 18 months of age.

The numbers and percentages of patients who achieve the ability to hold head erect without support by the final visit will be presented by overall, including subgroups, age at the time of treatment (≤4 months, >4 months). The observed proportion achieving the ability to hold head erect without support during the observation period will be compared to zero using a 1-sided Exact Binomial Test, along with the corresponding 97.5% confidence intervals. To make computation of the p-value possible, the value of 0.1% will be used in place of a literal zero.

8.3.2. Proportion of Patients that Achieve Ability to Roll Over

The ability to roll from back to both sides is defined by BSIDv03 Gross Motor Subtest Item #20 and will be based upon the maximum function achieved at any time point up to 18 months of age.

The numbers and percentages of patients who achieve the ability to roll from back to both sides by the final visit will be presented by overall and by subgroups, including subgroups, age at the time of treatment (≤4 months, >4 months). The observed proportion attaining the ability to roll from back to both sides during the observation period will be compared to zero using a 1-sided Exact Binomial Test, along with the corresponding 97.5% confidence intervals. To make computation of the p-value possible, the value of 0.1% will be used in place of a literal zero.

8.3.3. Proportion of Patients that Achieve Functional Independent Sitting for at least 30 seconds

The ability to sit without support is defined by BSIDv03 item #26 and will be based upon the maximum function achieved at any time point up to 18 months of age.

The numbers and percentages of patients who achieve the ability to sit with support will be presented by overall and by subgroups, including age at the time of treatment (≤4 months, >4 months). The observed proportion attaining the ability to sit with support during the observation period will be compared to zero using a 1-sided Exact Binomial Test, along with the corresponding 97.5% confidence intervals. To make computation of the p-value possible, the value of 0.1% will be used in place of a literal zero.

8.3.4. Proportion of Patients that Achieve ability to stand with assistance for at least 2 seconds

The ability to stand with assistance for at least 2 seconds is defined by BSIDv03 Gross Motor Subtest Item #33 and will be based upon the maximum function achieved at any time point up to 18 months of age.

The numbers and percentages of patients who achieve the ability to stand with assistance for at least 2 seconds will be presented by overall and by subgroups, including age at the time of treatment (≤4 months, >4 months). The observed proportion attaining the ability to stand with assistance for at least 2 seconds during the observation period will be compared to zero using a 1-sided Exact Binomial Test, along with the corresponding 97.5% confidence intervals. To make computation of the p-value possible, the value of 0.1% will be used in place of a literal zero.

8.3.5. Proportion of Patients that Achieve ability to crawl forwards on hands and knees for a distance of at least 5 feet

The ability ability to crawl forwards on hands and knees for a distance of at least 5 feet is defined by BSIDv03 Gross Motor Subtest Item #34 and will be based upon the maximum function achieved at any time point up to 18 months of age.

The numbers and percentages of patients who ability to crawl forwards on hands and knees for a distance of at least 5 feet during will be presented by overall and by subgroups, including age at the time of treatment (≤4 months, >4 months). The observed proportion attaining the ability to crawl forwards on hands and knees for a distance of at least 5 feet during the observation period will be compared to zero using a 1-sided Exact Binomial Test, along with the corresponding 97.5% confidence intervals. To make computation of the p-value possible, the value of 0.1% will be used in place of a literal zero.

8.3.6. Proportion of Patients that Achieve ability to pull to stand

The ability ability to pull to stand is defined by BSIDv03 Gross Motor Subtest Item #35 and will be based upon the maximum function achieved at any time point up to 18 months of age.

The numbers and percentages of patients who achieve the ability to pull to stand will be presented by overall and by subgroups, including age at the time of treatment (≤4 months, >4 months). The observed proportion attaining the ability to pull to stand during the observation period will be compared to zero using a 1-sided Exact Binomial Test, along with the corresponding 97.5% confidence intervals. To make computation of the p-value possible, the value of 0.1% will be used in place of a literal zero.

8.3.7. Proportion of Patients that Achieve ability to walk with assistance

The ability ability to walk with assistance is defined by BSIDv03 Gross Motor Subtest Item #37 and will be based upon the maximum function achieved at any time point up to 18 months of age.

The numbers and percentages of patients who achieve the ability to walk with assistance will be presented by overall and by subgroups, including age at the time of treatment (≤4 months, >4 months). The observed proportion attaining the ability to walk with assistance during the observation period will be compared to zero using a 1-sided Exact Binomial Test, along with the corresponding 97.5% confidence intervals. To make computation of the p-value possible, the value of 0.1% will be used in place of a literal zero.

8.3.8. Proportion of Patients that Achieve ability to stand alone for at least 3 seconds

The ability ability to stand alone for at least 3 seconds is defined by BSIDv03 Gross Motor Subtest Item #40 and will be based upon the maximum function achieved at any time point up to 18 months of age.

The numbers and percentages of patients who achieve the ability to stand alone for at least 3 seconds will be presented by overall and by subgroups, including age at the time of treatment (≤4 months, >4 months). The observed proportion attaining the ability to stand alone for at least 3 seconds during the observation period will be compared to zero using a 1-sided Exact Binomial Test, along with the corresponding 97.5% confidence intervals. To make computation of the p-value possible, the value of 0.1% will be used in place of a literal zero.

8.3.9. Proportion of Patients that Achieve ability to walk at least 3 steps without support

The ability ability to walk at least 3 steps without support is defined by BSIDv03 Gross Motor Subtest Item #42 and will be based upon the maximum function achieved at any time point up to 18 months of age.

The numbers and percentages of patients who achieve the ability to walk at least 3 steps without support will be presented by overall and by subgroups, including age at the time of treatment (≤4 months, >4 months). The observed proportion attaining the ability to walk at least 3 steps without support during the observation period will be compared to zero using a 1-sided Exact Binomial Test, along with the corresponding 97.5% confidence intervals. To make computation of the p-value possible, the value of 0.1% will be used in place of a literal zero.

8.3.10. Proportion of Patients That Achieve Ability to Crawl

The ability to crawl is defined by WHO development milestones and will be based upon the maximum function achieved at any time point up to 18 months of age.

The numbers and percentages of patients who achieve the ability to crawl will be presented by overall and by subgroups, including subgroups, age at the time of treatment (≤4 months, >4 months). The observed proportion attaining the ability to crawl during the observation period will be compared to zero using a 1-sided Exact Binomial Test, along with the corresponding 97.5% confidence intervals. To make computation of the p-value possible, the value of 0.1% will be used in place of a literal zero.

8.3.11. Proportion of Patients that Achieve the Ability to Stand with Assistance

The ability to stand with assistance is defined by WHO development milestones and will be based upon the maximum function achieved at any time point up to 18 months of age.

The numbers and percentages of patients who achieve the ability to stand with assistance will be presented by overall and by subgroups, including subgroups, age at the time of treatment (≤4 months, >4 months). The observed proportion attaining the ability to stand with assistance during the observation period will be compared to zero using a 1-sided Exact Binomial Test, along with the corresponding 97.5% confidence intervals. To make computation of the p-value possible, the value of 0.1% will be used in place of a literal zero.

8.3.12. Proportion of Patients that Achieve Ability to Stand Alone

The ability to stand alone is defined by WHO development milestones and will be based upon the maximum function achieved at any time point up to 18 months of age.

The numbers and percentages of patients who achieve ability to stand alone by the final visit will be presented by overall and by subgroups, including subgroups, age at the time of treatment (≤4 months, >4 months). The observed proportion achieving ability to stand alone during the observation period will be compared to zero using a 1-sided Exact Binomial Test, along with the corresponding 97.5% confidence intervals. To make computation of the p-value possible, the value of 0.1% will be used in place of a literal zero.

8.3.13. Proportion of Patients that Achieve Ability to Walk with Assistance

The ability to walk with assistance is defined by WHO development milestones and will be based upon the maximum function achieved at any time point up to 18 months of age.

The numbers and percentages of patients who achieve ability to walk with assistance by the final visit will be presented by overall and by subgroups, including subgroups, age at the time of treatment (≤4 months, >4 months). The observed proportion achieving ability to walk with assistance during the observation period will be compared to zero using a 1-sided Exact Binomial Test, along with the corresponding 97.5% confidence intervals. To make computation of the p-value possible, the value of 0.1% will be used in place of a literal zero.

8.3.14. Proportion of Patients that Achieve Ability to Walk Alone

The ability to walk alone is defined by WHO development milestones and will be based upon the maximum function achieved at any time point up to 18 months of age.

The numbers and percentages of patients who achieve ability to walk alone by the final visit will be presented by overall and by subgroups, including subgroups, age at the time of treatment (≤4 months, >4 months). The observed proportion achieving ability to walk alone during the observation period will be compared to zero using a 1-sided Exact Binomial Test, along with the corresponding 97.5% confidence intervals. To make computation of the p-value possible, the value of 0.1% will be used in place of a literal zero.

8.3.15. Improvement of Raw Score and Scaled Score using Bayley Scales of Infant and Toddler Development (Version 3), Fine and Gross Motor Function Subtests

Descriptive statistics for change from baseline in Fine and Gross motor function raw and scaled scores will be presented at each Study Day based visit, including mean, standard deviation, median, minimum, and maximum values.

Descriptive statistics for change from baseline in Fine and Gross motor function raw and scaled scores will also be presented at each Age Based visit, including mean, standard deviation, median, minimum, and maximum values.

The change from baseline of Bayley raw score to maximum score at any visit will be presented at by overall. An ANCOVA model with subgroup variable, as the factor and baseline Bayley score as a covariate, will be used to evaluate the comparability of the maximum change of Bayley score between subgroups.

The change from baseline will be analyzed by using mixed model with repeated measurement (MMRM). The model will include the change from baseline as the dependent variable, and fixed effects of subgroup, visit, and a covariate of baseline, age at baseline, and interactions of subgroup*visit, baseline*visit. A Toeplitz covariance structure will be assumed initially to model the within-patient errors; however, if the covariance structure results in non-convergence, the structure of compound symmetry will be used. The mean change from baseline (%), 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 through Month 18.

8.3.16. Maximum change from baseline of CHOP-INTEND score

The change from baseline of CHOP-INTEND score to maximum score at any visit up to 18 months of age will be presented at by overall and by subgroups, including subgroups, age at the time of treatment (≤4 months, >4 months). An ANCOVA model with subgroup variable as the factor and baseline CHOP-INTEND score as a covariate, will be used to evaluate the comparability of the maximum change of CHOP-INTEND score between subgroups.

8.3.17. Achievement of Selected Threshold Scores on CHOP-INTEND

8.3.17.1. Proportion of Patients Achieving CHOP-INTEND Score ≥40

The numbers and percentages of patients who achieve CHOP-INTEND score ≥40 by the final visit will be presented by overall and by subgroups, including subgroups, age at the time of treatment (≤4 months, >4 months). The observed proportion achieving CHOP-INTEND score ≥40 during the observation period will be compared to zero using a 1-sided Exact Binomial Test, along with the corresponding 97.5% confidence intervals. To make computation of the p-value possible, the value of 0.1% will be used in place of a literal zero. Patients with baseline scores ≥40 will not be considered in the proportional analysis.

8.3.17.2. Proportion of Patients Achieving CHOP-INTEND Score ≥50

The numbers and percentages of patients who achieve CHOPINTEND score ≥50 by the final visit will be presented by overall and by subgroups, including subgroups, age at the time of treatment (≤4 months, >4 months). The observed proportion achieving CHOP-INTEND score ≥50 during the observation period will be compared to zero using a 1-sided Exact Binomial Test, along with the corresponding 97.5% confidence intervals. To make computation of the p-

value possible, the value of 0.1% will be used in place of a literal zero. Patients with baseline scores ≥50 will not be considered in the proportional analysis.

8.3.17.3. Proportion of Patients Achieving CHOP-INTEND Score ≥58

The numbers and percentages of patients who achieve CHOP-INTEND score ≥58 by the final visit will be presented by overall and by subgroups, including subgroups, age at the time of treatment (≤4 months, >4 months). The observed proportion achieving CHOP-INTEND score ≥58 during the observation period will be compared to zero using a 1-sided Exact Binomial Test, along with the corresponding 97.5% confidence intervals. To make computation of the p-value possible, the value of 0.1% will be used in place of a literal zero. Patients with baseline scores ≥58 will not be considered in the proportional analysis.

8.3.17.4. Proportion of patients with the ability to remain independent of ventilator support

The numbers and percentages of patients with the ability to remain independent of ventilator support at 18 months of age will be presented by overall and by subgroups, including subgroups, age at the time of treatment (≤4 months, >4 months).

The observed proportion achieving independent of ventilator support at 18 months of age will be compared to zero using a 1-sided Exact Binomial Test, along with the corresponding 97.5% confidence intervals. To make computation of the p-value possible, the value of 0.1% will be used in place of a literal zero.

8.3.18. Proportion of patients maintaining the Ability to Thrive

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

- the ability to tolerate thin liquids as demonstrated through a formal swallowing test
- does not receive nutrition through mechanical support (i.e., feeding tube)
- maintains weight (>third percentile for age and gender as defined by WHO guidelines) at 18 months of age.

A patient who is not required mechanical support to receive nutrition at the date of 18 months of age visit is considered to meet the feeding criterion.

The observed proportion of patients maintaining the ability to thrive at 18 months will be compared to zero using a 1-sided Exact Binomial Test, along with the corresponding 97.5% confidence intervals. To make computation of the p-value possible, the value of 0.1% will be used in place of a literal zero.

Additionally, the proportion of patients meeting each of the criteria will be summarized in ability to thrive population.

8.3.19. Age of Achievement of Independent Sitting (for 10 seconds)

The ability to sit independently will be defined as by WHO MGRS. Descriptive statistics will be presented for these endpoints, for patients that achieve functional independent sitting, including mean, standard deviation, median, minimum, and maximum values. Milestone of Independent Sitting collected outside of SAP-defined visit window of Month 18 of age visit will be included in the analysis.

8.4. Multiplicity

The primary and secondary efficacy hypotheses are to be tested in a hierarchy approach that specifies the order in which they are to be evaluated. This strategy requires that only when the primary endpoint meets significance will the secondary endpoint be tested. The order of testing will be: first, independent sitting (primary endpoint); second, event -free survival (secondary endpoint). Such pre-specification of the order of testing within hierarchical framework strongly prevents type 1 error inflation due to multiplicity.

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, as well as duration of infusion, whether the entire volume was delivered, and whether the infusion was interrupted.

10.2. Treatment Compliance

Weight-adjusted dose, in vg/kg, of AVXS-101 administered and dosing compliance of AVXS-101 will be summarized.

The treatment compliance of each patient at Baseline will calculated as:

Percentage of compliance (%) = 100% * total volume administered / planned dose

The treatment compliance will be summarized by actual product received.

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 the last study visit. 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 using Safety Set.

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 patients 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
    - Any AE reported as Possibly, Probably, or Definitely Related will be consolidated into a single category and summarized as “Related” in the tables

- Any AE reported as Unlikely Related or Unrelated will be consolidated into a single category and summarized as “Not Related” in the tables

- Grade 3 and 4 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 the final analysis tables (i.e., after database lock of the study), just the two relatedness categories (“Related”, “Not Related”) will be presented.

For an interim run of this table (i.e., using an interim data cut), it is possible that there will be AEs with unknown relationship present. If this situation occurs, the table will present a third row (“Unknown”) in addition to “Related” and “Not Related” and the AEs with unknown relationship will be summarized in the “Unknown” row. If there are no AEs with unknown relationship in the interim cut of data, the “Unknown” row in the table will not be included.

Data listings, patient narratives, etc. will present the relationship to study treatment as collected on the CRF.

10.3.2.2. Adverse Event by SOC and PT

The following summaries of adverse events will be generated:

- Serious adverse events for screen failure patients only;
- Treatment-emergent adverse events;
- Treatment-emergent adverse events occurring in two (2) or more patients;
- Treatment-emergent adverse events categorized as "related" to AVXS-101;
- Serious treatment-emergent adverse events;
- Serious treatment-emergent adverse events related to AVXS-101;
- Grade 4 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").
- Treatment-emergent adverse events of special interest (see section 10.3.2.4);

For all adverse event summaries, the number of treatment-emergent adverse events, the number and percentage of patients experiencing treatment-emergent adverse events will be tabulated according to SOC and PT. Patients 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). Patients reporting more than one adverse event within a

SOC will be counted only once for that SOC. Patients reporting more than one adverse event will be counted only once in the overall total.

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

10.3.2.3. Adverse Event by PT

The number of treatment-emergent adverse events and the number and percentage of patients experiencing treatment-emergent adverse events will be tabulated according to preferred term and sorted by overall frequency. Similar summaries will be provided for Grade 3 and 4 treatment-emergent adverse events and treatment-emergent adverse events with a "possibly related" to AVXS-101 categorization.

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:

- Hepatotoxicity, identified via the following SMQs:
  - Hepatic failure, fibrosis and cirrhosis and other liver damage-related conditions (SMQ)
  - Hepatic disorders (SMQ)
- Thrombocytopenia, identified via the following SMQs/HLT:
  - Hematopoietic thrombocytopenia (SMQ)
  - Haemorrhages (SMQ)
  - All TEAEs which code into the “Platelet disorders NEC” HLT
- Cardiac events, identified via the following SMQs:
  - Ischemic heart disease (SMQ)
  - Cardiomyopathy (SMQ)
  - Cardiac arrhythmias (SMQ)
  - Embolic and thrombotic events (SMQ)
  - Myocardial infarction (SMQ)
- Ganglionitis, with potential cases identified by reviewing all TEAEs which code to the “Nervous System Disorders” SOC

For each AE of interest category, the number and percentage of patients experiencing at least one TEAE in the search for the event of interest will be presented overall. AEs of special interest will be summarized by SOC and PT overall.

10.3.2.5. Related Adverse Events by Maximum Grade

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 grade of each preferred term. If a patient has an adverse event with

unknown grade, then the patient will be counted in the grade category of "unknown," even if the patient has another occurrence of the same event with a grade present.

10.3.2.6. Adverse Events by Maximum Severity Grade Level

Treatment-emergent adverse events will be summarized by maximum CTCAE 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 patient has a non-serious adverse event with unknown grade, then the patient will be counted in the severity grade level category of "unknown," even if the patient has another occurrence of the same event with a grade 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 patient has an adverse event with unknown relationship, then the patient will be counted in the relationship category of "unknown," even if the patient has another occurrence of the same event with a relationship present. The only exception is if the patient has another occurrence of the same adverse event with a relationship assessment of "possibly related" or higher. In this case, the patient will be counted under the "related" category.

10.4. Laboratory Evaluations

10.4.1. Analysis of Laboratory Data

Safety laboratory data and genetic diagnosis laboratory data generated by Q2, the designated central laboratory, will be used in all analyses for genetic diagnosis. Also, safety laboratory data generated by local lab will be used for safety analysis. Immunoassay data generated by CTL will be used in all analyses.

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), and red cell distribution width (RDW).

Chemistry variables include: albumin, alanine aminotransferase (ALT/SGPT), alkaline phosphatase, aspartate aminotransferase (AST/SGOT), blood urea nitrogen (BUN), creatinine, gamma glutamyl transferase (GGT), glucose, serum total bilirubin, direct bilirubin, total creatine kinase (CK), CK-MB, and electrolytes.

Urinalysis variables include: specific gravity, pH, ketones, glucose, protein, blood, leukocyte esterase, nitrites, bilirubin, 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, binding antibody titer to SMN, T-cell response to AAV9 and SMN.

The Criteria for PCS Laboratory Findings are maintained outside of this SAP.

10.4.3. Statistical Methods

Clinical laboratory test will be summarized 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. Natural history control data is not available for this evaluation.

Mean changes from baseline to End of Study/Early Termination 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 will be summarized. 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 Period.

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 (as defined by the central laboratory and based on CTCAE v4.03) 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 patients 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. 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 patients with hemoglobin abnormalities.

For patients 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.

The number and percentage of patients meeting the following criteria will be summarized overall.

- ALT ≥ 3x ULN with bilirubin < 2x ULN
- AST ≥ 3x ULN with bilirubin < 2x ULN
- ALT ≥ 5x ULN with bilirubin < 2x ULN
- AST ≥ 5x ULN with bilirubin < 2x ULN
- ALT ≥ 10x ULN with bilirubin < 2x ULN

- AST ≥ 10x ULN with bilirubin < 2x ULN
- ALT ≥ 20x ULN with bilirubin < 2x ULN
- AST ≥ 20x ULN with bilirubin < 2x ULN
- ALT ≥ 3x ULN with bilirubin ≥ 2x ULN
- AST ≥ 3x ULN with bilirubin ≥ 2x ULN
- ALT ≥ 5x ULN with bilirubin ≥ 2x ULN
- AST ≥ 5x ULN with bilirubin ≥ 2x ULN
- ALT ≥ 10x ULN with bilirubin ≥ 2x ULN
- AST ≥ 10x ULN with bilirubin ≥ 2x ULN
- ALT and AST ≥ 3x ULN with bilirubin < 2x ULN
- ALT and AST ≥ 5x ULN with bilirubin < 2x ULN
- ALT and AST ≥ 10x ULN with bilirubin < 2x ULN
- ALT and AST ≥ 3x ULN with bilirubin ≥ 2x ULN
- ALT and AST ≥ 5x ULN with bilirubin ≥ 2x ULN
- ALT and AST ≥ 10x ULN with bilirubin ≥ 2x ULN
- ALT and AST ≥ 20x ULN with bilirubin ≥ 2x 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). For patients meeting any elevation criterion, a corresponding listing of all ALT, AST, alkaline phosphatase, and total, direct, and indirect bilirubin values will be provided.

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 two components:

1. Subjects showing ALT or AST elevations, greater than 3xULN, combined with serum TBL to ≥2xULN, without initial findings of cholestasis (elevated serum ALP);
2. No other reason can be found to explain the combination of increased ALT or AST and TBL, such as viral hepatitis A, B, or C, pre-existing 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 on the safety set. In addition, vital signs results will be flagged as Potentially Clinically Significant (PCS) if they meet the pre-specified criteria. The number and percent of patients meeting each PCS criterion will be summarized starting at Day 1 and continuing through the end of study.

10.6. ECG

10.6.1. Additional Safety Endpoint: Cardiovascular Safety Evaluations

Serial ECG data will be collected at scheduled time points and provided in an external dataset. A 12-lead ECG will also be conducted at scheduled visits. The baseline value will be latest measure prior to dosing. Mean changes from baseline, standard deviation, minimum, maximum, and median to End of Study/Early Termination visit will be summarized for HR, PR, QR, QRS, QTcF.

The central reviewer will identify abnormal ECGs that are PCS, the definitions for which are maintained outside of this document.

A listing of all PCS ECGs will be provided. Summaries of ECG data will be presented overall.

10.7. 12-Lead Holter Monitor

A Holter monitor will continuously record the patient's 12-lead ECG for a total of 72 hours from Day –1 (24 hours prior to the start of gene replacement therapy infusion) through 48 hours after the start of infusion. On Days -1 to Day 3, serial ECG data will be pulled from the Holter monitor at time points of 'pre-dose', '2 hour', '4 hour', '6 hour', '8 hour', '12 hour', '24 hour',

'36 hour' and '48 hour' post-dose and assessed by a central reviewer. The summaries of Holter monitor data will be done by the actual treatment received overall.

10.8. 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 patient with the corresponding result on the baseline physical exam. Descriptive statistics only are planned for this data.

A listing of physical examination findings including 'Normal', 'Abnormal Not Clinically Significant', 'Abnormal Clinically Significant' and "Abnormal CS, related to underlying disease" will be provided. Additionally, "Abnormal CS, related to underlying disease" will be categorized into "Abnormal Clinically Significant" and presented in the shift table. The Shift table summarizing the shift of physical examination results ('Normal', 'Abnormal Not Clinically Significant', 'Abnormal Clinically Significant') for examination items from baseline to post-baseline visits will be presented. A table summarizing observed and change from baseline values for head circumference will be produced.

10.9. Pulmonary Exam

Patients will be assessed by a pulmonologist at the time points specified in the Schedule of Assessments on of the pulmonologist and/or investigator. Non-invasive ventilatory support equipment will be provided by AveXis, Inc. through a third-party vendor.

Pulmonary exam results will be presented in a listing.

10.10. Echocardiogram

Echocardiograms will be conducted at the time points specified in the Schedule of Assessments. Clinically significant, treatment-emergent findings (as determined locally) may be reported as AEs. A listing of echocardiogram results (abnormal/normal, etc.) and findings (left ventricle function, patent foramen ovale, or other) will be provided for all screening and post-baseline visits.

Additionally, echocardiogram data will be provided to an external cardiologist for centralized review. For any future analyses, the centrally reviewed data will be considered the primary echocardiogram source data.

Descriptive statistics for change from baseline in echocardiogram parameters, including but not limited to LV EF Single Plan 2D, will be presented at each visit, including mean, standard deviation, median, minimum, and maximum values.

10.11. Ventilatory support

The number of hours per day of non-invasive positive pressure ventilator in the intervals between each post-baseline visit will be summarized overall using descriptive statistics.

10.12. Other Safety

10.12.1. Additional Safety Endpoint: Immunologic Response

Immunoreactivity to AAV9 and hSMN will be monitored by the collection of samples at Screening, Day 7, Day 14, Day 21, and Day 30. 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 patients responding Positive or Negative for antibody response at each time point will be summarized. Furthermore, the distribution of patients by titer level will be summarized.

Secondly, T-cell response to AAV9 and SMN will be measured by the collection of samples at Day 7 and 30 and determined through the quantification of number of Spot Forming Cells (SFC) per million Peripheral Blood Mononuclear Cells (10^6PBMC). SFC values >100 are considered positive. T-cell response to AAV9 will be measured separately in two peptide pools. T-cell response to SMN will be measured in a single peptide pool.

The number of SFC/10^6 PBMC will be summarized at each sample time point. For each post-infusion time point, the number and percent of patients responding positively for T-cell response will be summarized.

10.12.2. Vector Shedding

Saliva, urine, and stool samples will be collected for vector shedding in accordance with the Schedule of Study Assessments, which includes 24 hours (Day 2) and 48 hours (Day 3) post-dose.

A listing of vector shedding will be provided.

11. HEALTH ECONOMICS

The analysis of healthcare resource utilization data may be performed upon the data availability.

12. INTERIM ANALYSES

Not applicable for this SAP.

13. CHANGE FROM ANALYSIS PLANNED IN PROTOCOL

The efficacy completers population definition in the 9 Jan 2018 version of the protocol reflects:

- All treated patients who reach 14 months of age, OR
- All treated patients who meet discontinuation criteria, discontinue the study due to an AE or death

However, for this analysis the efficacy completers analysis population will consist of:

- All treated patients who reach 14 months of age for the survival endpoint **or up to and including 18 months of age for the endpoint of achievement of functional independent sitting**, OR
- All treated patients who meet discontinuation criteria, discontinue the study due to an AE, or experience death

14. REFERENCE LIST

1 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.

2 Glanzman AM, McDermott MP, Montes J, et al. Validation of the Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders (CHOP INTEND). Pediatr Phys Ther. 2011;23(4):322-326.

3 Glanzman AM, Mazzone E, Main M, et al. The Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders (CHOP INTEND): test development and reliability. Neuromusc Disord. 2010;20(3):155-161.

4 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.

5 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.

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

7 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.

8 Wang CH, Finkel RS, Bertini ES, Schroth M, Simonds A, Wong B, Aloysius A, Morrison L, Main M, Crawford TO, Trela A; Participants of the International Conference on SMA Standard of Care. Consensus Statement for Standard of Care in Spinal Muscular Atrophy. J Child Neurol. 2007 Aug; 22(8): 1027-49.

15. PROGRAMMING CONSIDERATIONS

See TLF documentation for details.

16. QUALITY CONTROL

Refer to Quality Control Plan for SAS programs.

17. APPENDICES

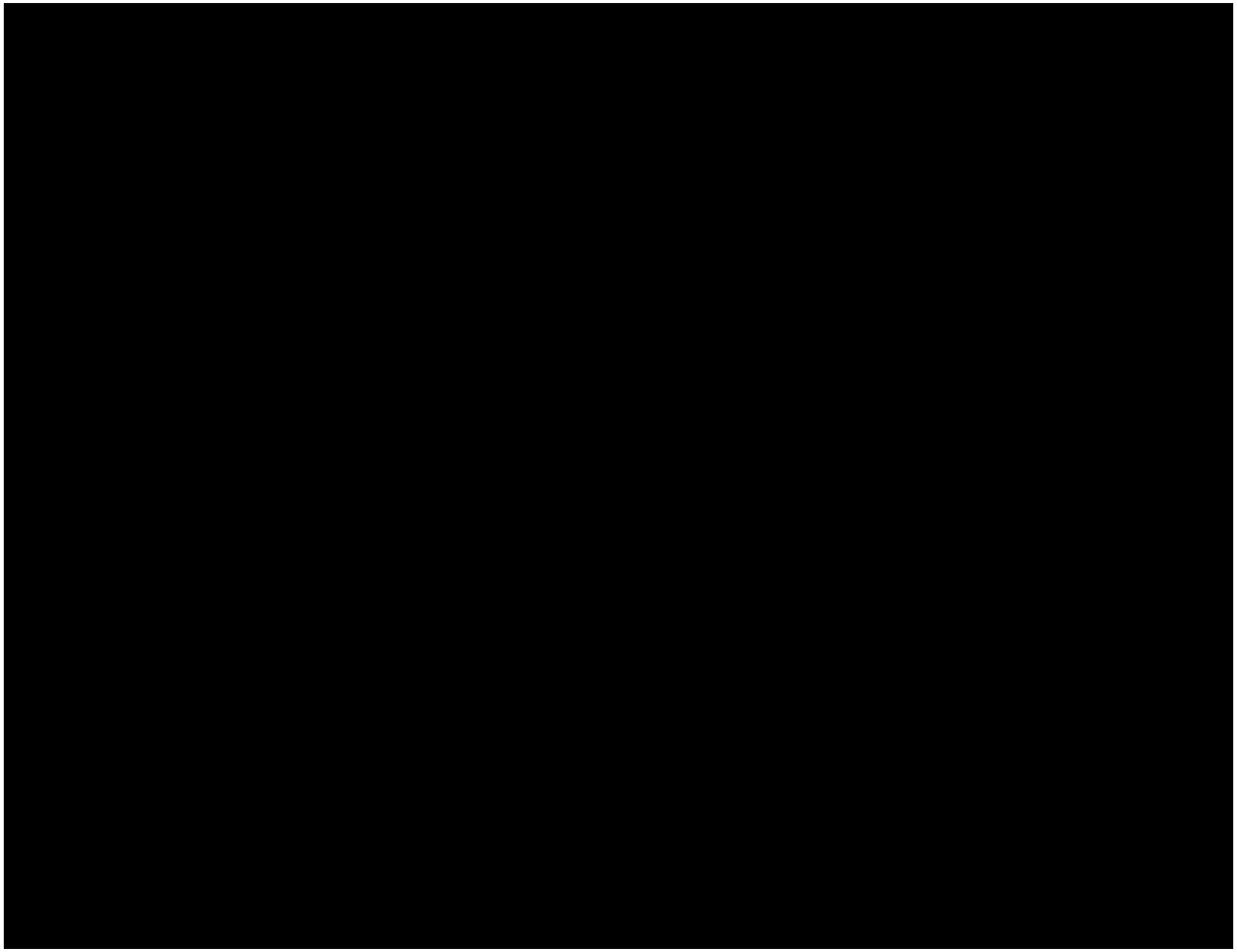
17.1. Performance Criteria for Bayley Scales of Infant and Toddler Development (Version 3) Developmental Milestones

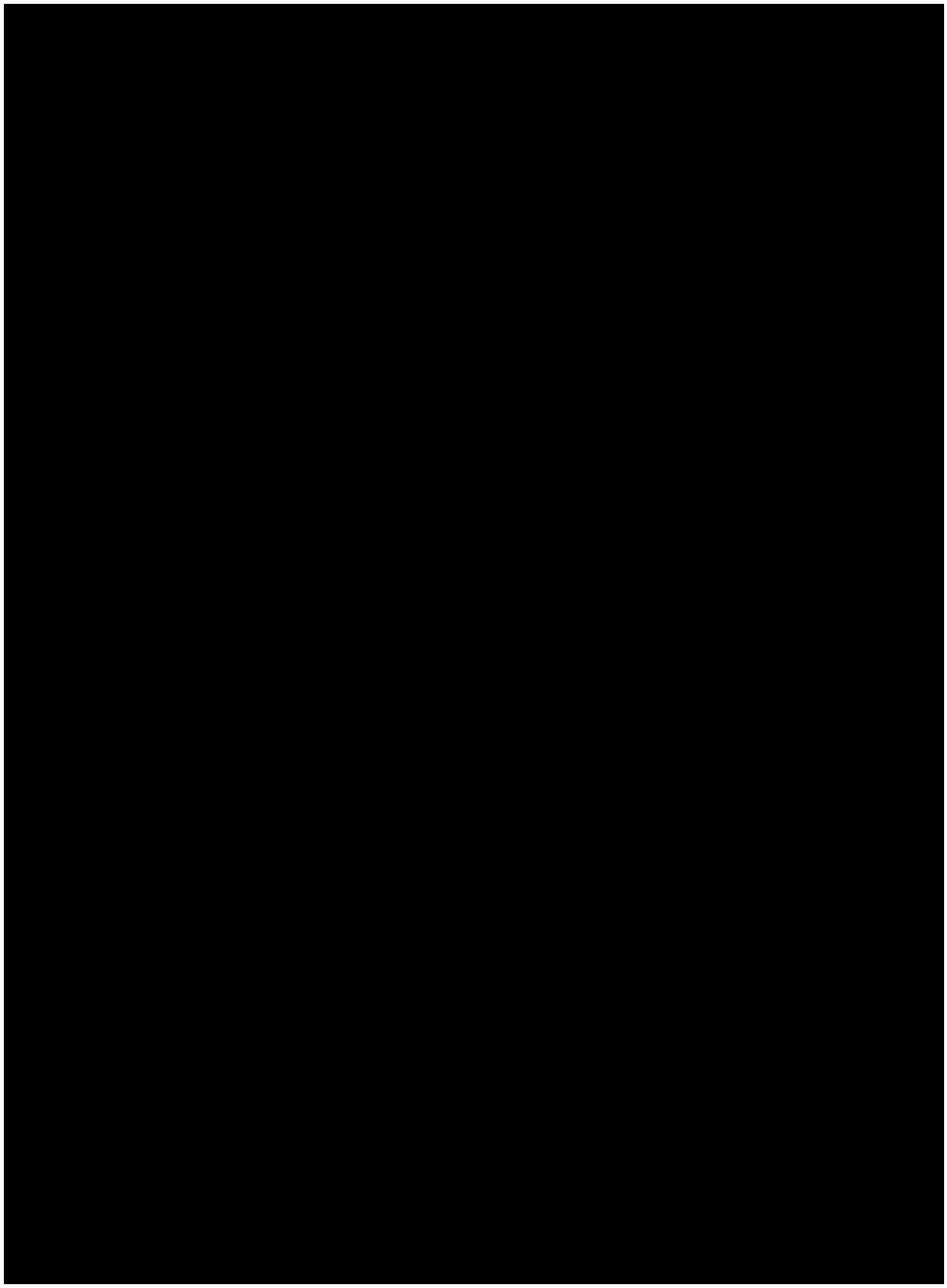
Developmental Milestone	[redacted]
Head Control – Gross Motor Subtest Item #4	
Rolls from Back to Sides – Gross Motor Subtest Item #20	
Sits Without Support – Gross Motor Subtest Item #26	
Stands with Assistance - Gross Motor Subtest Item #33	
Crawls – Gross Motor Subtest Item #34	
Pulls to Stand – Gross Motor Subtest Item #35	
Walks with Assistance – Gross Motor Subtest Item #37	
Stands Alone – Gross Motor Subtest Item #40	
Walks Alone – Gross Motor Subtest Item #42	

17.2. PERFORMANCE CRITERIA FOR WORLD HEALTH ORGANIZATION (WHO) DEVELOPMENTAL MILESTONES

Gross Motor Milestone	Performance Criteria
Sitting without support	[redacted]
Hands-and-knees crawling	
Standing with assistance	
Walking with assistance	
Standing alone	
Walking alone	

WHO = World Health Organization
Source: World Health Organization Multicentre Growth Reference Trial Group





17.4. SUMMARY OF CHANGES

The section below highlights content changes represented in version 2.0 of the SAP.

Global updates made through the entire document and not specified in the list of section-specific changes below:

Minor wording and formatting updates have been made throughout the SAP in order to correct typographical errors and to align wording across sections. These changes did not change the intent of the analysis.

The approval signature page was updated to follow the latest template.

Section 2.1 was updated to remove [redacted] as the responsible CRO for statistical and programming support.

Section 3.3 was updated to add new developmental milestones to be consistent with protocol version 7.0.

Section 3.4 was updated to revise safety objective to be consistent with protocol version 7.0.

Section 3.5 was updated to adopt the latest study design schematic figure.

Section 3.5 was updated to adopt the latest version of schedule of assessments.

Section 3.7 was updated to use 0.025 as significant level for one-sided test in primary objective.

Section 4.2 was updated to add new developmental milestones to be consistent with protocol version 7.0.

Section 4.3 was updated to revise the safety endpoints to be consistent with protocol version 7.0.

Section 4.6.7 was updated to revise cardiovascular safety evaluations due to protocol updates

Section 5.6 was updated to revise the survival rate at month 14 of age in PNCR data.

Section 5.9 was updated to replace AveXis SOP by ICH E3 Guidelines.

Section 6.2.3 was updated to redefine the final treatment values due to COVID pandemic cases.

Section 6.3 was updated to revise the strategy to hand missing values.

Section 6.4.1 was updated to revise the visit window due to Day 44 and Day 72 visit overlap.

Section 6.5 was updated to update the timing of data cutoff/database lock and corresponding statistical analyses.

Section 6.6 was updated to change the expected the number patients enrolled in a single site.

Section 6.7 was updated to simply the subgroups.

Section 7.4.1 was updated to add conversion to prophylactic prednisolone when other glucocorticosteroids was used.

Section 8.3 was updated to restrict analyses on ECP and revise the summary tables for data collected during COVID pandemic period.

Section 8.3.1-8.3.14 were updated to add new developmental milestones analyses.

Section 10.3.2.4 was updated to revise the categories of AESI.

Section 10.4.2 was updated to remove the PCS criteria, which will be versioned out of SAP.

Section 10.4.3 was updated to revise the criteria, which will be summarized

Section 10.5 was updated to remove the PCS criteria, which will be versioned out of SAP.

Section 10.6.1 was updated to revise the ECG summary information.

Section 10.7 was updated to add holter monitoring.

Section 10.12.1 was updated to revise the immunologic response.

Section 10.12.2 was updated to add vector shedding.