Protocol Number: AVXS-101-CL-304

Official Title: A Global Study of a Single, One-Time Dose of AVXS-101 Delivered to Infants with Genetically Diagnosed and Pre-symptomatic Spinal Muscular Atrophy with Multiple Copies of SMN2

NCT Number: NCT03505099

Document Date: 02-Jul-2021

STATISTICAL ANALYSIS PLAN

Protocol Number and Title:	AVXS-101-CL-304 A Global Study of a Single, One-Time Dose of AVXS-101 Delivered to Infants with Genetically Diagnosed and Pre-symptomatic Spinal Muscular Atrophy with Multiple Copies of *SMN2*
Protocol Version and Date:	6.0 / 28Jul2020 Japan-specific 5.0 / 08Sep2020
Author(s):	[redacted] Biostatistician, Contractor for AveXis Inc.
SAP Version:	3.0
SAP Version Date:	02 JULY 2021

REVISION HISTORY

Version #	Date (dd-Mmm-yyyy)	Document Owner	Revision Summary
0.1	21-Feb-2018	[redacted]	Initial release version
0.2	7-March-2018	[redacted]	Review comments on version 0.1 incorporated
0.3	29-March-2018	[redacted]	Review comments on version 0.2 incorporated
1.0	29-March-2018	[redacted]	SAP finalized and signed
1.1	28-Dec-2018	[redacted]	Initial release version of amendment 1
1.2	16-Feb-2019	[redacted]	Review comments on version 1.1 incorporated
1.3	09-Jan-2020	[redacted]	Updates throughout based on protocol amendment 4 (v5.0) and questions from programming team.
1.4	07-April-2020	[redacted]	Updates throughout based on questions from programming team.
1.5	25-Sep-2020	[redacted]	Updated throughout based on output review and protocol amendment.
1.6	14-Dec-2020	[redacted]	Updates throughout based on 2-copy cohort dry-run.
1.7	27-Jan-2021	[redacted]	Updates throughout based on 2-copy cohort dry-run.
2.0	27-Jan-2021	[redacted]	SAP v2.0 finalized
2.1	28-May-2021	[redacted]	Updates pertaining to 3-copy cohort throughout based on feedback from the 2-copy cohort final run.
3.0	02-July-2021	[redacted]	SAP v3.0 finalized

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

APPROVALS

[redacted]

Biostatistician Date (dd-Mmm-yyyy)
[redacted]
AveXis, Inc.

[redacted]

Senior Manager, Clinical Trial Operations Date (dd-Mmm-yyyy)
[redacted]
AveXis, Inc

[redacted]

Senior Medical Director, Clinical Development Date (dd-Mmm-yyyy)
[redacted]
AveXis, Inc

[redacted]

Medical Writing Date (dd-Mmm-yyyy)
[redacted]
AveXis, Inc

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. GLOSSARY OF ABBREVIATIONS

Abbreviation	Description
AAV / AAV9	Adeno Associated Virus 9
AE	Adverse Event
ALP	Alkaline Phosphatase
ALT	Alanine Aminotransferase
AST	Aspartate Aminotransferase
BiPAP	Bilevel Positive Airway Pressure
BSIDv03	Bayley Scales of Infant and Toddler Development version 3
BUN	Blood Urea Nitrogen
CI	Confidence Interval
CDISC	Clinical Data Interchange Standards Consortium
CHOP-INTEND	Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders
CRF/eCRF	Case Report Form / electronic Case Report Form
CTCAE	Common Terminology Criteria for Adverse Events
DILI	Drug-induced Liver Injury
DMC	Data Monitoring Committee
DSMB	Data Safety Monitoring Board
ECG	Electrocardiogram
EOS	End of Study
GGT	Gamma Glutamyl Transferase
HFMSE	Hammersmith Functional Motor Scale-Expanded
ICD-10	International Statistical Classification of Diseases and Related Health Problems
IFN-γ	Interferon Gamma
IRB	Institutional Review Board
ITT	Intent-to-treat

Abbreviation	Description
LFE	Liver Function Enzyme
LOCF	Last Observation Carried Forward
MedDRA	Medical Dictionary for Regulatory Activities
PCS	Potentially Clinically Significant
PNCR	Pediatric Neuromuscular Clinical Research
PT	Preferred Term
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD	Secure Digital
SMN	Survival Motor Neuron
SMN1	Survival Motor Neuron 1 gene
SMN2	Survival Motor Neuron 2 gene
SMQ	Standard MedDRA Query
SOC	System Organ Class
SOP	Standard Operating Procedure
TBL	Total Bilirubin
TEAE	Treatment Emergent Adverse Event
TLF	Table, Listing and Figure
ULN	Upper Limit Normal
WHO	World Health Organization
WHO-MGRS	World Health Organization-Multicentre Growth Reference Study

1.1. KEY DEFINITIONS

Term	Definition
AE	Adverse Event (AE): Any untoward medical occurrence in a clinical investigation patient. An AE 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 Baseline will be defined and displayed in days. Age at Baseline = (Dose date – date of birth +1).
Baseline	Baseline, unless otherwise specified in the Statistical Analysis Plan (SAP) sections, refers to the last measurement or evaluation made prior to the infusion of AVXS-101.
First Dose Date	The first dose date, Study Day 1, will be the date of administration of the study drug after enrollment.
MedDRA	Medical Dictionary for Regulatory Activities (MedDRA) is a medical terminology used to classify AE information associated with the use of biopharmaceuticals and other medical products.
Study Day	First dose of study medication will be administered on Study Day 1. There is no Study Day 0. If the event happened on or after the first dose date, Study Day is defined as event date – first dose date + 1. If the event happened prior to the first dose date, Study Day is defined as event date – first dose date.

2. PURPOSE

The purpose of this statistical analysis plan (SAP) is to ensure that the data listings, summary tables and figures that will be produced, and the statistical methodologies that will be used, are complete and appropriate to allow valid conclusions regarding the study objectives. The SAP was drafted with respect to the AVXS-101-CL-304 protocol version 6.0, dated 28 July 2020, and Japan-specific protocol version 4.0, dated 27 December 2019, entitled *A Global Study of a Single, One-Time Dose of AVXS-101 Delivered to Infants with Genetically Diagnosed and Pre-Symptomatic Spinal Muscular Atrophy with Multiple Copies of SMN2.*

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. (AveXis) is responsible for ownership and approval of the SAP.

A Contract Research Organization (CRO) selected by AveXis will derive the data sets according to Clinical Data Interchange Standards Consortium (CDISC) standards and create data set specifications based on the SAP. The CRO will perform the statistical analyses and is responsible for the production and quality control of all tables, figures and listings.

2.2. TIMING OF ANALYSES

The analysis for each Survival Motor Neuron 2 gene (*SMN2)* copy number cohort will be completed separately at such time that enrollment in the respective cohort is complete and the last patient has completed the End of Study (EOS) visit at the respective age or has discontinued study follow-up. Specifically:

- For *SMN2*=2, the EOS visit will occur at 18 months of age
- For *SMN2*=3, the EOS visit will occur at 24 months of age

3. STUDY OBJECTIVES

3.1. PRIMARY OBJECTIVES

3.1.1. Safety

The safety objectives are to:

- Evaluate the safety of AVXS-101 through incidence of adverse events (AEs) and/or serious adverse events (SAEs)
- Assess the safety of AVXS-101 based on the change from baseline in clinical laboratory parameters

3.1.2. Efficacy

Primary efficacy objectives are defined independently for each cohort.

3.1.2.1. Patients with bi-allelic Survival Motor Neuron 1 gene (SMN1) deletions and 2 copies of SMN2

The primary objective is to assess the efficacy of AVXS-101 by demonstrating [redacted] as defined by Bayley Scales of Infant and Toddler Development version 3 (BSIDv03) Gross Motor Subtest Item #26 at any visit up to 18 months of age.

3.1.2.2. Patients with bi-allelic SMN1 deletions and 3 copies of SMN2

The primary objective is to assess the efficacy of AVXS-101 based on the proportion of patients achieving the ability to stand alone for at least 3 seconds as defined by BSIDv03 Gross Motor Subtest Item #40 at any visit up to 24 months of age.

3.2. SECONDARY OBJECTIVES

3.2.1. Efficacy

Secondary efficacy objectives are defined independently for each cohort.

3.2.1.1. Patients with bi-allelic SMN1 deletions and 2 copies of SMN2

Secondary objectives include:

- Assess the efficacy of AVXS-101 based on survival, defined as avoidance of death or the requirement of permanent ventilation in the absence of acute illness or perioperatively as assessed at 14 months of age.
- Assess the efficacy of AVXS-101 by demonstrating the ability to maintain weight at or above the third percentile without need for non-oral/mechanical feeding support at any visit up to 18 months of age.

3.2.1.2. Patients with bi-allelic SMN1 deletions and 3 copies of SMN2

Assess the efficacy of AVXS-101 by demonstrating the ability to walk alone with coordination, defined as [redacted] as defined by BSIDv03 Gross Motor Subtest Item #43 at any visit up to 24 months of age.

3.3. EXPLORATORY OBJECTIVES

Exploratory efficacy objectives are defined independently for each cohort.

3.3.1. Efficacy

3.3.1.1. Patients with bi-allelic SMN1 deletions and 2 copies of SMN2

Exploratory objectives include:

[redacted]

3.3.1.2. Patients with bi-allelic SMN1 deletions and 3 copies of SMN2

Exploratory objectives include:

[redacted]

[redacted]

3.4. STATISTICAL HYPOTHESES

The primary and secondary efficacy analyses will be based on patients with bi-allelic deletion mutations of *SMN1*. Study cohorts will be based upon *SMN2* copy number, whereby patients

with 2 and 3 copies of *SMN2* will be subject to separate analyses. These analyses are to test the superiority of AVXS-101 to the results of natural observation studies using the historical data of Pediatric Neuromuscular Clinical Research (PNCR) network.

Two cohorts of patients (i.e., 2 copies of *SMN2* and 3 copies of *SMN2*) will be evaluated separately in this study. Family-wise type I error is controlled within each population cohort, as each cohort is considered its own experiment.

3.4.1. Hypothesis of Primary and Secondary Efficacy Endpoints for Patients with Bi-Allelic SMN1 Deletions and 2 Copies of SMN2

3.4.1.1. Primary Efficacy Endpoint

The primary efficacy endpoint for patients with bi-allelic *SMN1* deletions and 2 copies of *SMN2* is the proportion of patients who achieve the ability to [redacted] (as defined by BSIDv03 Gross Motor Subtest Item #26) at any visit up to the 18 months of age study visit.

The primary efficacy endpoint hypothesis for patients with bi-allelic *SMN1* deletions and 2 copies of *SMN2* is:

$$H_0: P_{AVXS\text{-}101} \leq 0.1\%$$

$$H_a: P_{AVXS\text{-}101} > 0.1\%$$

where *P* is the proportion of patients [redacted] at any visit up to the 18 months of age study visit. Based upon two widely cited natural history studies of the disease (NeuroNext [1], PNCR [2]), it is expected that no patient meeting the study entrance criteria (*SMN2* copy number of 2 without the *SMN2* gene modifier mutation [c.859G>C]) would be expected to attain the ability to sit without support. Due to computational considerations the comparison will be made to 0.1% in lieu of zero.

Testing will be performed using 1-sided exact test for a binomial proportion with $\alpha = 0.025$.

3.4.1.2. Secondary Efficacy Endpoints

There are two secondary efficacy endpoints defined for the cohort of patients with 2 copies of *SMN2*:

- Proportion of patients that have survived and have not required permanent ventilation in the absence of acute reversible illness and perioperatively, assessed at 14 months of age
- Proportion of patients that have achieved the ability to maintain weight at or above the 3rd percentile without need for non-oral/mechanical feeding support at any visit up to 18 months of age

The hypothesis for the first secondary efficacy endpoint for patients with bi-allelic *SMN1* deletions and 2 copies of *SMN2* to be tested is:

$$H_0: P_{AVXS\text{-}101} = P_{HISTORICAL\text{-}PNCR}$$

$$H_a: P_{AVXS\text{-}101} \neq P_{HISTORICAL\text{-}PNCR}$$

where *P* is the proportion surviving and not requiring ventilation at 14 months of age.

Formal testing for the proportion of patients surviving and not requiring permanent ventilation will only be conducted if the result for the primary endpoint (functional independent sitting) is statistically significant. Survival will be compared with the PNCR population-matched control cohort using a two sample 2-sided Fishers exact test with $\alpha = 0.05$. Only if the null hypothesis of equality in proportion of functional independent sitting is rejected at $p<0.025$, will the result of the test for this first secondary endpoint be considered statistically significant.

The natural history comparison dataset drawn from the PNCR for statistical testing of this secondary endpoint will include all patients with SMA Type 1, 2 copies of SMN2, age at SMA onset ≤ 6 months, and age at SMA diagnosis ≤ 2 years. The *SMN2* modifier mutation (c.859G>C) was not assessed in the PNCR study cohort.

Based upon this approach, patient-level data from a cohort of 23 patients drawn from the PNCR Network natural history study of SMA will serve as a "population-matched" control cohort. Of this cohort, 6/23 (26.1%) of the natural history cohort survived and did not require ventilation at 14 months of age.

The hypothesis for the second secondary efficacy endpoint for patients with bi-allelic *SMN1* deletions and 2 copies of *SMN2* is:

$$H_0: P_{AVXS\text{-}101} \leq 0.1\%$$

$$H_a: P_{AVXS\text{-}101} > 0.1\%$$

where *P* is the proportion of patients who have achieved the ability to maintain weight at or above the third percentile without need for non-oral/mechanical feeding support at any visit up to 18 months of age. For comparison, the number of patients who maintained the ability to maintain weight at or above the third percentile and/or were independent of ventilatory support at 18 months of age in the PNCR database was essentially zero [5]. Due to computational considerations, the comparison will be made to 0.1% in lieu of zero.

Only if the primary endpoint (functional independent sitting) and the first secondary endpoint (survival) for patients with 2 copies of *SMN2* as described above meet statistical significance will this second secondary endpoint of ability to maintain weight at or above the third percentile without need for non-oral/mechanical feeding support be formally tested. Testing for this endpoint will be performed using 1-sided exact binomial test with $\alpha = 0.025$. Only if the null

hypothesis of equality in proportion of survival is rejected at $p<0.05$, will the result of the test for this second secondary endpoint be considered statistically significant.

This hierarchical approach strongly protects against Type I error within the 2-copy cohort.

3.4.2. Hypothesis of Primary and Secondary Efficacy Endpoints for Patients with Bi-Allelic SMN1 Deletions and 3 Copies of SMN2

3.4.2.1. Primary Efficacy Endpoint

The primary efficacy endpoint for patients with 3 copies of *SMN2* is the proportion of patients who achieve the ability to stand alone for at least 3 seconds (as defined by BSIDv03 Gross Motor Subtest Item #40) at any visit up to the 24 months of age study visit.

The primary efficacy endpoint hypothesis for patients with bi-allelic *SMN1* deletions and 3 copies of *SMN2* to be tested is:

$$H_0: P_{AVXS\text{-}101} = P_{HISTORICAL\text{-}PNCR}$$

$$H_a: P_{AVXS\text{-}101} \neq P_{HISTORICAL\text{-}PNCR}$$

where *P* is the proportion of patients who achieve the ability to stand alone for at least 3 seconds at any visit up to the 24 months of age study visit.

Testing will be performed using a 2-sided two sample Fisher exact test with $\alpha = 0.05$.

The natural history comparison dataset drawn from the PNCR will include all patients with SMA of any type who provided sufficient records for evaluation and met the basic entry criteria for this cohort of the AVXS101CL-304 study (3 copies of *SMN2*) and had at least one study visit with a non-missing assessment for Hammersmith Functional Motor Scale – Expanded (HFMSE) item #19. The *SMN2* modifier mutation (c.859G>C) was not assessed in the PNCR study cohort.

Based upon this approach, patient-level data from a cohort of 81 patients drawn from the PNCR Network natural history study of SMA will serve as a "population-matched" control cohort. This comparison cohort encompasses all 81 patients enrolled in the PNCR study who met the criteria of having 3 copies of *SMN2* and had at least one study visit with a non-missing assessment for Hammersmith item #19. Of this cohort, 19/81 (23.5%) of the natural history cohort attained the ability to stand alone (defined as achieving a score of 2 on item #19 of the HFMSE). Note that the achievement of a score of 2 on item #19 is not restricted to occurrences up to 24 months of age in the natural history cohort. Achievement of this score at any time while in the PNCR study will be considered.

3.4.2.2. Secondary Efficacy Endpoint

There is one secondary efficacy endpoint defined for the cohort of patients with 3 copies of *SMN2*:

Proportion of patients demonstrating the ability to walk alone with coordination, defined as [redacted] at any visit up to 24 months of age (as defined by BSIDv03 Gross Motor Subtest Item #43). The secondary efficacy endpoint hypothesis for patients with bi-allelic *SMN1* deletions and 3 copies of *SMN2* to be tested is:

$$H_0: P_{AVXS\text{-}101} = P_{HISTORICAL\text{-}PNCR}$$

$$H_a: P_{AVXS\text{-}101} \neq P_{HISTORICAL\text{-}PNCR}$$

where *P* is the proportion of patients who achieve the ability to walk alone with coordination at any visit up to the 24 months of age study visit.

Only if the primary endpoint (stand alone) meets statistical significance will this secondary endpoint be formally tested. Testing will be performed using a two sample 2-sided Fisher exact test with $\alpha = 0.05$.

The natural history comparison dataset drawn from the PNCR will include all patients with SMA of any type who provided sufficient records for evaluation and met the basic entry criteria for this cohort of the AVXS-101-CL-304 study (3 copies of *SMN2*) and had at least one study visit with a non-missing assessment for HFMSE item #20. The *SMN2* modifier mutation (c.859G>C) was not assessed in the PNCR study cohort.

Based upon this approach, patient-level data from a cohort of 81 patients drawn from the PNCR Network natural history study of SMA will serve as a "population-matched" control cohort. This comparison cohort encompasses all 81 patients enrolled in the PNCR study who met the criteria of having 3 copies of *SMN2* and had at least one study visit with a non-missing assessment for HFMSE item #20. Of the cohort that serves as the comparator in testing the secondary endpoint, 17/81 (21%) attained the ability to walk alone with coordination (defined as achieving a score of 2 on item #20 of the HFMSE). Note that the achievement of a score of 2 on item #20 is not restricted to occurrences up to 24 months of age in the natural history cohort. Achievement of this score at any time while in the PNCR study will be considered.

This hierarchical approach strongly protects against Type I error within the 3-copy cohort.

3.5. STUDY DESIGN

This is a Phase 3, open-label, single-arm study of a single, one-time dose of AVXS-101 (gene replacement therapy) in patients with SMA who meet enrollment criteria and are genetically defined by bi-allelic deletion of *SMN1* with 2 or 3 copies of survival motor neuron 2 gene (*SMN2*). Patients with *SMN1* point mutations or the *SMN2* gene modifier mutation (c.859G>C) may enroll in the study but will not be included in the efficacy analysis sets.

At least 14 patients with bi-allelic deletion of *SMN1* and 2 copies of *SMN2* and at least 12 patients with bi-allelic deletion of *SMN1* and 3 copies of *SMN2* that are ≤6 weeks of age at the time of gene replacement therapy (Day 1) will be enrolled.

The study includes a screening period, a gene replacement therapy period, and a follow-up period. 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 study enrollment. Patients who meet all of the entry criteria and do not meet any of the exclusion criteria will enter the in-patient gene replacement therapy period (Day –1 to Day 2). On Day –1, patients will be admitted to the hospital for pre-treatment baseline procedures. On Day 1, patients will receive a single, one-time intravenous (IV) infusion of AVXS-101 and will undergo in-patient safety monitoring over the next 24 hours. Patients may be discharged 24 hours after the infusion, based on Investigator judgment. During the outpatient follow-up period (Days 3 to EOS at 18 or 24 months of age, depending upon the respective *SMN2* copy number), patients will return at regularly scheduled intervals for efficacy and safety assessments until the EOS when the patient reaches 18 months of age (*SMN2* = 2) or 24 months of age (*SMN2* = 3). After the EOS visit, eligible patients will be asked to rollover into a long-term follow-up study.

Following dosing, follow-up visits will be conducted every week for the first four weeks. Subsequent visits will occur at Day 44, Month 2, Day 72, and Month 3, and then every 3 months, based on patient age, at 6, 9, 12, 15, and 18 months old and, if applicable, at 21 and 24 months old.

In an attempt to dampen the host immune response to the adeno-associated virus (AAV) derived therapy, all patients will receive prophylactic prednisolone at approximately 2 mg/kg/day (or an equivalent dose of another glucocorticoid if prednisolone is unavailable or in the opinion of the investigator prednisolone is not tolerated) on Day -1, Day 1, and Day 2, and then 1 mg/kg/day starting on Day 3 and until at least 30 days post-AVXS-101 infusion. After 30 days post-AVXS-101 infusion, the dose of prednisolone can be tapered for patients whose gamma glutamyl transferase (GGT), alanine aminotransferase (ALT) values, and aspartate aminotransferase (AST) values are below the threshold of 2 x Upper Limit of Normal (ULN) in accordance with the following treatment guideline: taper from 1 mg/kg/day to 0.5 mg/kg/day during Weeks 5 and 6 post-AVXS-101 infusion, then taper to 0.25 mg/kg/day during Weeks 7 and 8, and then discontinue prednisolone at Week 9.

For patients with 2 copies of SMN2 (SMN2=2), efficacy will be assessed by achievement of the key developmental milestone of functional independent sitting for at least 30 seconds as defined by BSIDv03 Gross Motor Subtest Item #26 at any visit up to 18 months of age. For patients with 3 copies of SMN2 (SMN2=3), efficacy will be assessed by achievement of the ability to stand without support for at least three seconds as defined by BSIDv03 Gross Motor Subtest Item #40 at any visit up to 24 months of age. Additional developmental milestones will be assessed using the WHO-MGRS and BSIDv03. Safety will be assessed through monitoring AEs, concomitant medication usage, physical examinations, vital sign assessments, cardiac assessments, and laboratory evaluations.

The primary efficacy analysis for each *SMN2* copy number cohort will be completed separately at such time that all patients in the cohort have either discontinued or completed the study.

A Data Safety Monitoring Board (DSMB)/Data Monitoring Committee (DMC) will review safety data on a quarterly basis. Refer to Section 11 of this SAP for additional details regarding the DSMB/DMC.

3.6. PATIENT SELECTION

Patients with pre-symptomatic genetically diagnosed SMA who are ≤6 weeks (≤42 days) of age at the time of gene replacement therapy (Day 1) with documented absence of the *SMN1* gene and 2 or 3 copies of the *SMN2* gene will be enrolled in this study. Patients may be of any gender and any racial or ethnic background.

3.6.1. Inclusion Criteria

See protocol for details.

3.6.2. Exclusion Criteria

See protocol for details.

3.7. DETERMINATION OF SAMPLE SIZE

At least 14 patients with bi-allelic *SMN1* deletions and 2 copies of *SMN2* and at least 12 patients with bi-allelic *SMN1* deletions and 3 copies of *SMN2* will be enrolled.

3.7.1. Patients with bi-allelic SMN1 deletions and 2 copies of SMN2

The primary efficacy endpoint for patients with bi-allelic *SMN1* deletions and 2 copies of *SMN2* is the proportion of patients who achieve the ability to sit without support for at least 30 seconds (as defined by BSIDv03 Gross Motor Subtest Item #26) at any visit up to the 18 months of age study visit.

Based upon two widely cited natural history studies of the disease (PNCR and NeuroNext), it is expected that no patient meeting the study entrance criteria (*SMN2* copy number of 2 without the *SMN2* gene modifier mutation [c.859G>C]) would be expected to attain the ability to sit without support. Based upon data from the completed AVXS-101-CL-101 study, at least 60% of treated patients with 2 copies of *SMN2* are expected to achieve the ability to sit without support for at least 30 seconds. With this degree of efficacy, a sample size of 14 patients would provide power of >90% to detect a significant difference compared with a rate of 0.1% (in lieu of zero) with $\alpha = 0.025$ using a 1-sided exact test for a binomial proportion.

3.7.2. Patients with bi-allelic SMN1 deletions and 3 copies of SMN2

The primary efficacy endpoint for patients with bi-allelic *SMN1* deletions and 3 copies of *SMN2* is the proportion of patients who achieve the ability to stand alone (as defined by BSIDv03 Gross Motor Subtest Item #40) for at least 3 seconds at any visit up to the 24 months of age study visit.

In the PNCR dataset, 19/81 (23.5%) of patients with *SMN2*=3 achieved the ability to stand alone. Extrapolating from the experience from the AVXS-101 Phase 1 study in infants with SMA Type 1, 85% of treated patients with bi-allelic *SMN1* deletions and 3 copies of *SMN2* are expected to achieve the ability to stand alone. With this degree of efficacy, a sample size of 12 patients would provide power of >90% to detect a significant difference compared with the population-matched control cohort (from the PNCR database) with $\alpha = 0.05$ using a two sample 2-sided Fisher exact test.

3.8. TREATMENT ASSIGNMENT AND BLINDING

The clinical study is a Phase 3, open-label, single-dose administration study. The treatment assignment will be conducted as specified in Section 3.5 of the SAP. No blinding will be conducted.

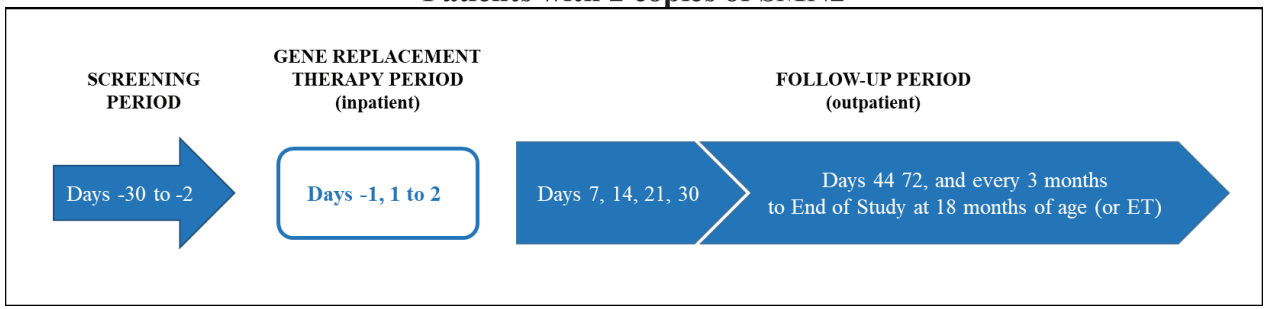
3.9. ADMINISTRATION OF STUDY MEDICATION

Refer to Section 3.5 of the SAP.

3.10. STUDY FLOWCHART

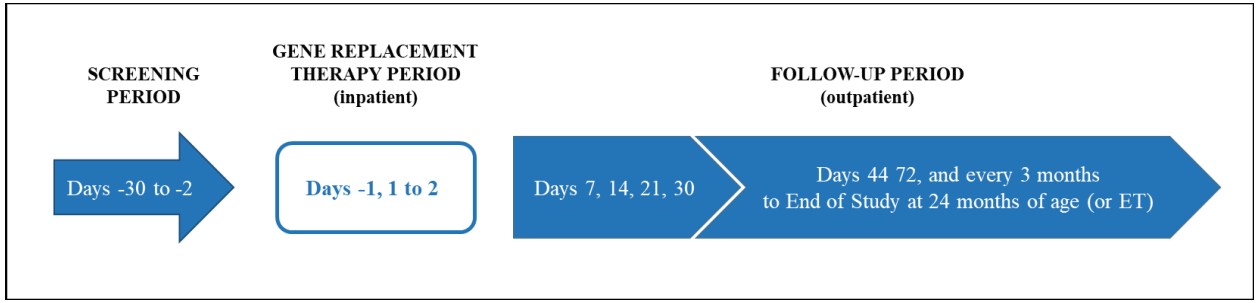
Figure 1: Study Design

Patients with 2 copies of SMN2



Note: After the EOS visit at 18 months of age, patients will be monitored for safety for 30 days after the EOS visit. Additionally, patients will be invited to participate in a long-term follow-up study conducted under a separate protocol. EOS = End of Study; ET = early termination

Patients with 3 copies of SMN2



Note: After the EOS visit at 24 months of age, patients will be monitored for safety for 30 days after the EOS visit. Additionally, patients will be will be invited to participate in a long-term follow-up study conducted under a separate protocol. EOS = End of Study; ET = early termination

3.11. PATIENTS WITH 4 COPIES OF SMN2

This study was initially designed to enroll at least 17 patients with bi-allelic deletion of SMN1 and 4 copies of *SMN2* that were ≤6 weeks of age at the time of gene replacement therapy. The primary objective of the 4-copy cohort was to assess the efficacy of AVXS-101 by demonstrating the ability to achieve a scaled score on the BSIDv03 Gross and Fine Motor Subtests within 1.5 standard deviations of chronological development reference standard as assessed at 36 months of age. The exploratory objective of the 4-copy cohort was to [redacted]

- [redacted]
- [redacted]

The protocol was subsequently amended (Protocol version 2 [amendment 1], dated 27-September-2018) to remove the cohort of patients with 4 copies of SMN2. Due to observations in toxicology studies being conducted at the time, it was decided to not include the 4-copy cohort until the preclinical safety of AVXS-101 was further investigated and better understood.

One patient was enrolled into this study around the same time as protocol version 2 was finalized. This patient was enrolled into the 3-copy *SMN2* cohort, but it was subsequently determined via genetic reconfirmation testing that the patient had 4 copies of *SMN2*.

Operationally, this situation was handled as follows. This patient was considered to have 4 copies of *SMN2*, based on the genetic confirmation testing. This 4-copy patient underwent all of the study assessments which were specified in the protocol for the 3-copy cohort, and was followed through 24 months of age. All data collected for this 4-copy patient have been summarized in a stand-alone set of data listings, as described throughout this SAP.

4. ENDPOINTS

4.1. PRIMARY EFFICACY ENDPOINTS

Primary efficacy endpoints will be assessed independently for each cohort. Assessment of developmental motor milestones will be primarily determined by the Site Clinical Evaluator (Physical Therapist, or Occupational Therapist) at the investigative sites. The primary source for motor milestone data will be based on documented video evidence captured by the Site Clinical Evaluator (Physical Therapist, or Occupational Therapist) or patient families. 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, Central Reviewer for concordance. Only developmental milestones that the Central Reviewer has confirmed or has additionally identified will be included for the primary analysis.

4.1.1. Patients with bi-allelic SMN1 deletions and 2 copies of SMN2

- Proportion of patients achieving the development milestone of functional independent [redacted] (as defined by BSIDv03 Gross Motor Subtest Item #26) at any visit up to 18 months of age.

4.1.2. Patients with bi-allelic SMN1 deletions and 3 copies of SMN2

- Proportion of patients achieving the ability to stand alone for at least three seconds (as defined by BSIDv03 Gross Motor Subtest Item #40) at any visit up to 24 months of age.

4.2. SECONDARY EFFICACY ENDPOINTS

Secondary efficacy endpoints will be assessed independently for each cohort.

4.2.1. Patients with bi-allelic SMN1 deletions and 2 copies of SMN2

- Proportion of patients that have survived and have not required permanent ventilation in the absence of acute reversible illness or perioperatively, assessed at 14 months of age.

- Proportion of patients that have achieved the ability to maintain weight at or above the third percentile without need for non-oral/mechanical feeding support at any visit up to 18 months of age

4.2.2. Patients with bi-allelic SMN1 deletions and 3 copies of SMN2

- Proportion of patients demonstrating the ability to walk alone with coordination, defined as [redacted] (as defined by BSIDv03 Gross Motor Subtest Item #43) at any visit up to 24 months of age.

4.3. EXPLORATORY EFFICACY ENDPOINTS

Exploratory efficacy endpoints will be assessed independently for each cohort.

4.3.1. Patients with bi-allelic SMN1 deletions and 2 copies of SMN2

- [redacted]
  - [redacted]
  - [redacted]
  - [redacted]
  - [redacted]

[redacted]

4.3.2. Patients with bi-allelic SMN1 deletions and 3 copies of SMN2

[redacted]

[redacted]

4.4. OTHER EFFICACY ENDPOINTS

4.4.1. Compound Motor Action Potential (CMAP)

Peroneal nerve CMAP amplitude will be measured by a qualified electrophysiologist using the procedures described in the CMAP manual. CMAP will be measured at Screening, every 6 months starting at 6 months of age, and EOS when the patient reaches 18 or 24 months of age (or early termination), depending on the respective cohort. The CMAP data (surface temperature, amplitude [mV] and area [mVmsec]) will be collected for centralized review and interpretation.

4.5. SAFETY ENDPOINTS

Safety endpoints will be assessed separately for each cohort.

4.5.1. Adverse Events

All AEs that occur from the start of gene replacement therapy infusion through the last study visit will be collected and recorded in the eCRF. All AEs reported will be assessed for their seriousness, relatedness to study treatment, relationship to study discontinuation, and severity

according to CTCAE version 4.03 criteria [3]. AEs will be coded using the most current version of MedDRA available at the time of database lock for each cohort.

4.5.2. Vital Signs, Body Weight and Length

Vital signs including blood pressure, heart rate, respiratory rate, pulse, temperature, and pulse oximetry as well as body weight and length will be collected at each study visit.

4.5.3. Physical Examination

Physical examinations will be conducted at each study visit as indicated in the study protocol schedule of assessments. Examinations will include review of the following systems: head, eyes, ears, nose and throat, lungs/thorax, cardiovascular, abdomen, musculoskeletal, dermatologic, lymphatic, neurologic and genitourinary.

The head circumference shall be measured with each physical examination. To measure head circumference, the examiner should securely wrap a flexible measuring tape around the circumference of the head, above the eyebrows over the broadest part of the forehead, above the ears, and over the most prominent part of the occiput. The measurement should be taken 3 times, and the largest measurement should be recorded to an accuracy of 0.1 centimeters.

4.5.4. Pulmonary Examination

Pulmonary examinations will be performed by a pulmonologist or appropriate individual as per standard institutional practice at each scheduled visit except Day 1.

During the study patients may be fitted with ventilatory support at the discretion of the pulmonologist (or appropriate individual as per standard institutional practice and/or Investigator). Non-invasive ventilatory support equipment will be provided by AveXis, Inc. through a third-party vendor if not covered by the patient's insurance. Should the patient require non-invasive ventilatory support at any time during the study, the usage should be recorded in the eCRF. Patients requiring non-invasive ventilatory support will be asked to bring their machine(s) to each study visit.

4.5.5. Laboratory Evaluations

Blood samples will be collected at each scheduled visit as specified in in the protocol schedule of assessments and measured for standard hematology and blood chemistry tests. Urine samples will be collected for standard urinalysis parameters.

4.5.6. Swallowing Test

A swallowing test will be performed at screening, every 6 months starting at 6 months of age, and at EOS to determine if the patient has signs of aspiration. If the test is positive for aspiration, the parent or guardian of the patient will be recommended to use an alternative method to oral

feeding. Once implemented, 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.

4.5.7. Cardiac Testing

4.5.7.1. Electrocardiogram (ECG)

Standard 12-lead ECGs will be performed at screening, Day -1, pre-dose on Day 1, Day 2, at the 3, 6, 9, and 12 months of age visits, and every six months thereafter. Data collected will include heart rate, as well as RR, PR, QRS, QT, and QTcF intervals. 12-lead ECG tracings will be interpreted locally; however, the tracings will be submitted to a centralized interpreter for inclusion in the study dataset.

4.5.7.2. Echocardiogram

An echocardiogram will be performed at screening, at the 3, 6, 9, and 12 months of age visits, and every six months thereafter. Echocardiograms will be interpreted locally; however, they will also be submitted to a centralized interpreter for inclusion in the study dataset.

4.5.7.3. Twenty-Four Hour Holter Monitoring

Twenty-four hour Holter monitoring will be performed at screening and Days -1 through 2 (while in hospital). Twenty-four hour Holter monitoring will also be performed at the 3, 6, 9, and 12 months of age visits, and every 6 months thereafter.

4.5.8. Photographs of infusion site

Photographs will be taken of the infusion site at each scheduled visit from Day 1 (post-dose) through Day 30 to monitor healing of the infusion site.

4.5.9. Prior and 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 World Health Organization (WHO) Drug dictionary. Medications will be coded using the most current version of the WHO Drug dictionary available at the time of database lock for each cohort.

4.6. PHARMACOKINETIC AND PHARMACODYNAMIC ENDPOINTS

Not applicable to this SAP.

4.7. HEALTH-ECONOMICS ENDPOINTS

Hospitalization information from time of birth will be captured in the eCRF.

5. ANALYSIS POPULATIONS

5.1. ALL ENROLLED POPULATION

The All Enrolled Population will consist of all patients enrolled (i.e., completed the informed consent process) who receive AVXS-101, including patients with *SMN1* point mutations and patients who are positive for the *SMN2* gene modifier mutation (c.859G>C). Unless specified otherwise, this population will be used for patient listings and for summaries of patient disposition.

5.2. SAFETY POPULATION

The Safety Population will include all patients who are given an AVXS-101 injection, including patients with *SMN1* point mutations and patients who are positive for the *SMN2* gene modifier mutation (c.859G>C).

Patients who are enrolled into the 2 or 3 copy *SMN2* cohort who are subsequently determined via the reconfirmation testing to have 4 copies of *SMN2* will still be included in the Safety Population, but their safety data will not be summarized along with all other patients in the cohort they were initially enrolled into. Safety data for patients who were determined to have 4 copies of SMN2 will be provided in separate, stand-alone listings.

The Safety Population will be used for all analyses of safety endpoints and for the presentation of patients in patient listings containing safety data.

5.3. INTENT TO TREAT (ITT) POPULATION

The Intent-to-Treat (ITT) Population will consist of all enrolled patients with bi-allelic *SMN1* deletions and 2 or 3 copies of *SMN2* without the *SMN2* gene modifier mutation (c.859G>C) who receive AVXS-101. Patients who are positive for the *SMN2* gene modifier mutation (c.859G>C) are excluded from the ITT Population.

Patients who are enrolled into the 2 or 3 copy *SMN2* cohort who are subsequently determined via the reconfirmation testing to have 4 copies of *SMN2* will be excluded from the ITT Population and their efficacy data will be listed separately. Patients will be analyzed according to the *SMN2* copy number determined via the reconfirmation testing. Only genetic results from the central laboratory or determined from local labs as indicated by the SMA diagnosis eCRF (for patients with indeterminant results from the central lab or if no central laboratory results are available) will be used to determine a patient's SMN2 copy number and inclusion in the ITT Population.

All efficacy analyses will be conducted using the ITT Population as the primary population.

5.4. EFFICACY COMPLETERS (EC) POPULATION

The Efficacy Completers (EC) Population, a subset of the ITT Population, will consist of all patients who complete the study. Patients who terminate early will not be included in the EC Population. The EOS visit can be conducted either remotely or at the study site.

5.5. NATURAL HISTORY STUDY POPULATION

One natural history study population is available for comparison with the study population. Distinct control populations drawn from the PNCR study [2] will be used as a comparison for the primary endpoint of sitting without support for at least 30 seconds (as defined by BSIDv03 Gross Motor Subtest Item #26) and the secondary endpoints of survival and the ability to maintain weight at or above the 3rd percentile without need for non-oral/mechanical feeding support at any visit up to 18 months of age for the cohort of patients with bi-allelic *SMN1* deletions and 2 copies of *SMN2* and for the primary and secondary endpoints (the ability to stand alone for at least 3 seconds as defined by BSIDv03 Gross Motor Subtest Item #40 and the ability to walk alone with coordination as defined by BSIDv03 Gross Motor Subtest Item #43) for the cohort of patients with bi-allelic *SMN1* deletions and 3 copies of *SMN2*. Exploratory endpoints will utilize the PNCR (2 or 3 copies of *SMN2*) cohorts as appropriate.

5.5.1. PNCR Population-Matched Control Population

The PNCR Natural History dataset [2] is drawn from a natural history study performed on a cohort of 337 patients with any form of SMA followed at 3 large, internationally recognized tertiary medical centers with significant expertise in the management of SMA [redacted]. 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 published in 2007 were used as a basis for providing uniform care among the study sites.

5.6. ALL SCREENED POPULATION

The All Screened Population will consist of all patients screened. Unless specified otherwise, this population will be used for the patient listing of screen failures.

5.7. PROTOCOL DEVIATIONS

Protocol deviations identified by the site or the study monitor are to be reported to the Institutional Review Board (IRB) according to the IRB's reporting guidelines. All deviations are to be recorded in the [redacted] database. They will be categorized in accordance with AveXis' Standard Operating Procedures (SOPs).

6. GENERAL ASPECTS FOR STATISTICAL ANALYSIS

6.1. GENERAL METHODS

In general, descriptive statistical methods will be used to summarize the data from this study. Hypothesis testing will be performed for the primary and secondary efficacy endpoints based upon the predetermined sample sizes. Unless stated otherwise, the term "descriptive statistics" refers to number of patients (n), mean, median, standard deviation, minimum, and maximum for continuous data and frequencies and percentages for categorical data. Unless noted otherwise, the data presented in listings will be sorted by patient number, and then by date within each patient number.

Efficacy analyses will be conducted on the ITT Population. In the event that the EC Population includes a subset of patients in the ITT Population, then select (primary and secondary) efficacy analyses will also be conducted on the EC Population. If the EC Population includes the same set of patients as are included in the ITT Population, then no efficacy analyses will be conducted on the EC Population. All safety analyses will be summarized using the Safety Population.

Efficacy and safety data from each cohort (patients with 2 and 3 copies of *SMN2*) will be summarized in separate sets of tables, listings, and figures. Additionally, efficacy data collected from patients with the *SMN2* gene modifier mutation (c.859G>C) will be listed in a separate sets of outputs, based on the number of *SMN2* copies these patients have (2 or 3). Efficacy and safety data for patients who were determined to have 4 copies of *SMN2* will be provided in listing format only.

The efficacy analyses will compare the primary and secondary efficacy endpoints of patients treated with AVXS-101 either directly to patients drawn from a widely (peer-reviewed) published natural history dataset collected by the PNCR network [2], or to assumed rates based on such information. For exploratory efficacy analyses, patients treated with AVXS-101 will be compared to patients drawn from the PNCR network natural history study where possible. For patients with the *SMN2* gene modifier mutation (c.859G>C), efficacy data will be summarized in listing format only.

All hypothesis testing will be conducted one-sided with significance level α=0.025 or two-sided with significance level α=0.05, as indicated.

All statistical analyses will be conducted with the SAS® software package version 9.3 or higher.

6.2. KEY DEFINITIONS

6.2.1. Baseline Values

Unless otherwise specified, the Baseline values of efficacy and safety parameters of the patients enrolled in the study are defined as the last non-missing measurement or assessment prior to the

dose of study medication. If multiple measurements are recorded on the same day, the last measurement recorded prior to the dose of study medication will be used as the baseline value.

The Baseline values of parameters of patients from the PNCR cohort are defined as those captured for the study-defined Baseline visit.

6.2.2. Study Day

For patients enrolled in the study, the only dose of study medication will be administered on Study Day 1. The study day for an event occurring on or after the dose date is defined as event date – dose date + 1. The study day for an event occurring prior to the dose date is defined as event date – dose date.

For patients from the PNCR dataset, the study day for an event occurring on or after the date of enrollment is defined as event date – enrollment date + 1. The study day for an event occurring prior to the date of enrollment is defined as event date – enrollment date.

6.2.3. Dose Date

The dose date (Study Day 1) is defined as the date of administration of study product.

6.2.4. Final Treatment Value

The final treatment value for each patient enrolled in the study is the last non-missing measurement collected after Study Day 1.

For the patients from the PNCR dataset, the final treatment value is the last non-missing measurement collected after Study Day 1 and on or prior to Visit Month 18 (2-copies of *SMN2* cohort) or Visit Month 24 (3-copies of *SMN2* cohort).

6.2.5. Final Study Value

The final study value for each patient enrolled in the study is the last non-missing measurement collected.

For patients from the PNCR dataset, the final study value is the measurement collected on or prior to Visit Month 18 (2-copies of *SMN2* cohort) or Visit Month 24 (3-copies of *SMN2* cohort).

6.2.6. Age at Baseline

Age at baseline for patients enrolled in the study is defined as the age at the dose date. Age at baseline for patients from the PNCR dataset is defined as age at the baseline visit. Age will be defined/displayed in days and/or months, depending on the data being summarized.

6.2.7. Determination of SMN Copy Number

A blood sample will be collected at the screening visit for reconfirmation of SMN1 deletions/mutations, *SMN2* copy number, and absence of exon 7 gene modifier mutation (c.859G>C). Only genetic results from the central laboratory or determined from local labs as indicated on the SMA diagnosis eCRF (for patients with indeterminant results from the central lab or if no central laboratory results are available) will be used to determine a patient's SMN2 copy number and inclusion in the ITT Population.

If a patient is enrolled in either the 2 copy or 3 copy *SMN2* cohort and is subsequently determined (via the reconfirmation testing) to have 4 copies of *SMN2*, then the efficacy and safety results for that patient will be summarized separately in data listings and will not be summarized together with the other patients in the cohort they were originally enrolled into.

If the genetic test results are indeterminate, then the patient will be considered to have the higher copy number indicated in the result. For example, a genetic result of "Atypical (between 2-3 copies)" would result in a patient being assigned to the 3 copy cohort.

6.2.8. Age at Study Visit/Event

The patient age at which a study visit or event occurs is defined as study visit date (or event date) – date of birth + 1. The study day for an event or study visit occurring prior to the dose date is defined as study visit date (or event date) – date of birth.

6.3. MISSING DATA

For responder-type endpoints (such as achievement of motor milestones), missing values will be imputed as non-responsive. For survival-type endpoints (such as time to respiratory intervention), patients who prematurely withdraw from the study will be censored at the time of withdrawal). For the endpoint of survival at 14 months of age for patients in Cohort 1, only patients who prematurely withdrew from the study prior to turning 14 months of age will need to be censored at the time of withdrawal; patients who prematurely withdrew after turning 14 months of age would not require censoring for this endpoint.

For the CHOP-INTEND instrument, rules suggested by the producers of these assessments will be followed in calculating scores when individual question/items may be missing. If these rules do not provide enough information for calculating a score, then the endpoint will be considered as having a missing value.

There are 16 items in CHOP-INTEND assessment. A rating of Brazelton behavioral states was recorded for each item. As suggested by the manual, the optimal state for testing is state 4 ("alert, with bright look") and 5 ("eyes open"). If a patient 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 items 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. Similarly, any CHOP-INTEND assessments which have been flagged as "invalid" on the eCRF will not be included in any analyses.

Generally, missing values for safety endpoints will not be imputed. Imputation of missing or incomplete dates will only be performed on AEs and concomitant medications for determining the timing relative to the dose of study product unless otherwise specified. The rules of imputation of missing dates are specified in Table 1 and Table 2. Partial or missing dates will be listed as recorded in the eCRF. No imputed date will be earlier than the patients' birth date.

Table 1: Rules of Date Imputation: Pre-Dose Assessments

Partial Missing Start or Stop Date	Imputed Start 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
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
Missing month, but year and day are present	Missing month imputed as January

Table 2: Rules of Date Imputation: Post-Dose Assessments

Partial Missing Start or Stop Date	Imputed Start Date
Missing month and day, but the year is present	Date of dose
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
Missing month, but year and day are present	Month of dose

6.4. DERIVED AND TRANSFORMED DATA

6.4.1. Primary Efficacy Variables

BSIDv03 is a standardized, norm-referenced infant assessment. The Gross and Fine Motor subtest alone will be completed on Study Day 30, Month 2, Month 3, every 3 months starting at 6 months of age, and at EOS when the patient reaches 18 or 24 months of age (or early termination), dependent upon the cohort the patient is in. No missing scores will be imputed.

The achievement of significant motor milestones will be assessed by the site clinical evaluator using a standard Motor Milestone Development Survey shown in Table 3 with definitions of

each milestone driven by the Bayley Scales. Milestones will be assessed by the physical therapist at the investigational site; the physical assessments will be captured on video. Video evidence of milestone achievement will then be reviewed and confirmed by an external independent Central Reviewer.

Table 3: Motor Milestone Development Survey – Bayley Scales

Developmental Milestone- Bayley Scales Item Number	**Performance Criteria**
Head Control – Gross Motor Subtest Item #4	[redacted]
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 #43	

6.4.1.1. Patients with bi-allelic SMN1 deletions and 2 copies of SMN2

The primary efficacy endpoint for patients with bi-allelic *SMN1* deletions and 2 copies of *SMN2* is the proportion of patients achieving [redacted] at any visit up to the 18-months of age study visit. The ability to sit independently for at least 30 seconds will be evaluated by the site clinical evaluator on BSIDv03 Gross Motor Subtest will be classified as responders. Otherwise, the patient will be classified as a non-responder.

6.4.1.2. Patients with bi-allelic SMN1 deletions and 3 copies of SMN2

The primary efficacy endpoint for patients with bi-allelic *SMN1* deletions and 3 copies of *SMN2* is the proportion of patients achieving the ability to stand alone for at least 3 seconds at any visit up to the 24 months of age study visit. The ability to stand alone will be evaluated by the site clinical evaluator on BSIDV03 Gross Motor Subtest item # 40 at visits specified by the Schedule of Assessments. Patients who achieve standing alone for at least 3 seconds at any post-procedure

visit up to and including the 24 months of age study visit (age ≤749 days) will be classified as responders. Otherwise, the patient will be classified as a non-responder.

Patient-level data from a cohort of 81 patients having 3 copies of *SMN2* will be drawn from the historical dataset of the PNCR network to serve as a population-matched control for the analysis of this endpoint. Patients from the PNCR database will be considered to have achieved the endpoint of ability to stand alone for at least 3 seconds if they have a score of 2 on item #19 of the HFMSE at any visit. The HFMSE definition for "stands alone for at least 3 seconds" is the same as the definition that is used for Bayley Scales item #40.

6.4.2. Secondary Efficacy Variables

6.4.2.1. Patients with bi-allelic SMN1 deletions and 2 copies of SMN2

The first secondary efficacy endpoint for patients with bi-allelic *SMN1* deletions and 2 copies of *SMN2* is the proportion of patients surviving and not requiring permanent ventilation at 14 months of age. This endpoint is defined as the avoidance of both (a) death and (b) permanent ventilation, which is defined by tracheostomy or by the requirement of ≥16 hours of respiratory assistance per day (via noninvasive ventilatory support) for ≥14 consecutive days in the absence of an acute reversible illness, excluding perioperative ventilation.

Non-invasive ventilatory support usage data will be recorded in the eCRF.

An "acute reversible illness" is defined as any condition other than SMA 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 (e.g., surgery) or medication (e.g., 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 patient received medications that could impair or interfere with respiratory function.

For a patient 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 patient 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 patient 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.

The presence/absence of an acute reversible illness, excluding perioperative ventilation, will be assessed by the investigator and documented on the "Ventilatory Support – Noninvasive" CRF.

Patient-level data from a cohort of 23 patients having 2 copies of *SMN2* will be drawn from the historical dataset of the PNCR network to serve as a population-matched control for the analysis of this endpoint. Among them, 17/23 (73.9%) of 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 13.6 month of age. The approximate survival rate at 13.6 months of age in the control cohort is 6/23 (26.1%).

Patients who terminate the study prior to reaching 14 months of age for any reason will be censored at the time of withdrawal.

The second secondary efficacy endpoint for patients with bi-allelic *SMN1* deletions and 2 copies of *SMN2* is the proportion of patients who have achieved the ability to maintain weight at or above the third percentile without need for non-oral/mechanical feeding support at any visit up to 18 months of age (age ≤569 days).

The ability to maintain weight at or above the third percentile without need for non-oral/mechanical feeding support is defined by meeting the following criteria at each visit up to 18 months of age:

- Does not receive nutrition through mechanical support (i.e., feeding tube)
- Maintains weight (≥third percentile for age and sex as defined by WHO guidelines) consistent with the patient's age at the assessment.

The third percentile weights for age and sex relevant for this study are as defined by the World Health Organization Child Growth Standards [6]. If the patient's recorded weight at every visit is at or above the third percentile for the patient's sex and age at the visit, and the patient never received nutrition through mechanical support, then they meet the definition of ability to maintain weight at or above the third percentile without need for non-oral/mechanical feeding support. If a patient ever received nutrition through mechanical support, and/or if the patient's weight drops below the third percentile for the patient's sex and age at any visit, then they do not meet the definition of ability to maintain weight at or above the third percentile without need for non-oral/mechanical feeding support.

For the purposes of determining whether a patient's recorded weight is at or above the third percentile, the patient's age in days at the visit will first be determined. Age will be converted to

both weeks (derived as age in days divided by 7) and months (derived as age in days divided by 30). Age in weeks and months will be rounded to the nearest whole week or month as follows:

For visits/assessments collected up to and including the Age 3 Months visit: Age in weeks will be rounded up to the next whole week for values ≥ X.50 and will be rounded down to the previous whole week for values < X.50. For example:

- An age of 2.6 weeks will be rounded up to 3 weeks; an age of 5.8 weeks will be rounded up to 6 weeks
- An age of 5.4 weeks will be rounded down to 5 weeks; an age of 0.4 weeks will be rounded down to 0 weeks

For visits/assessments collected after the Age 3 Months visit: Age in months will be rounded up to the next whole month for values ≥ X.50 and will be rounded down to the previous whole month for values < X.50. For example:

- An age of 3.6 months will be rounded up to 4 months; an age of 5.8 months will be rounded up to 6 months
- An age of 5.4 months will be rounded down to 5 months; an age of 12.2 months will be rounded down to 12 months

The corresponding third percentile weight for that (rounded) age and the patient's sex will be determined based on the World Health Organization (WHO) Child Growth Standards. For visits/assessments collected up to and including the Age 3 Months visit, the WHO Child Growth Standards *by week* for 0-13 weeks will be applied. For visits/assessments collected after the Age 3 Months visit, the WHO Child Growth Standards *by month* for 0-24 months will be applied.

For comparison, the number of patients who maintained the ability to thrive and/or were independent of ventilatory support at 18 months of age in the PNCR database was essentially zero [5].

An additional analysis using the definition of ability to thrive will be used for the 2-copy cohort whereby all three of the following criteria need to be met at the 18 months of age visit:

- Ability to tolerate thin liquids as demonstrated through a formal swallowing test, defined as follows:
    - Consistency tested is "Very thin" and the result is "Normal swallow", "Functional swallow", or "Safe for swallowing"; or
    - Consistency tested is "Thin" and the result is "Normal swallow", "Functional swallow", or "Safe for swallowing"
- Does not receive nutrition through mechanical support (i.e., feeding tube)
- Maintains weight (≥third percentile for age and sex as defined by WHO guidelines) consistent with the patient's age at the assessment.

With regards to the formal swallowing test, if there are multiple results available for the same consistency tested at the same time point, but with different results (i.e., 'Aspiration' and 'Normal swallow' for consistency tested 'Thin'), then a 'worst case' rule will be applied, whereby the worst result (in the example provided, 'Aspiration') will be considered to be the swallowing test result for the consistency tested at the time point.

If the patient's recorded weight at 18 months of age is at or above the third percentile, and the patient is not receiving nutrition through mechanical support at the 18 months of age visit, and the patient has evidence of being able to tolerate thin liquids as demonstrated through a formal swallowing test conducted at the 18 months of age visit, then they meet the definition of ability to thrive at that visit.

The number of patients who maintained the ability thrive and/or were independent of ventilatory support at 18 months of age in the PNCR database was essentially zero [2, 5].

6.4.2.2. Patients with bi-allelic SMN1 deletions and 3 copies of SMN2

The secondary efficacy endpoint for patients with bi-allelic *SMN1* deletions and 3 copies of *SMN2* is the proportion of patients who achieve the ability to walk alone with coordination at any visit up to the 24 months of age study visit. This endpoint is defined as [redacted] at any visit up to and including 24 months of age (age ≤749 days). The ability to walk alone will be evaluated by the site clinical evaluator on BSIDv03 Gross Motor Subtest Item # 43.

Patient-level data from a cohort of 81 patients having 3 copies of *SMN2* will be drawn from the historical dataset of the PNCR network to serve as a population-matched control for the analysis of this endpoint. Patients from the PNCR database will be considered to have achieved the endpoint of ability to walk alone if they have a score of 2 on item #20 of the HFMSE, defined as the ability to walk more than 4 steps unaided.

6.4.3. Exploratory Efficacy Variables

6.4.3.1. Patients with bi-allelic SMN1 deletions and 2 copies of SMN2

[redacted]

[redacted]

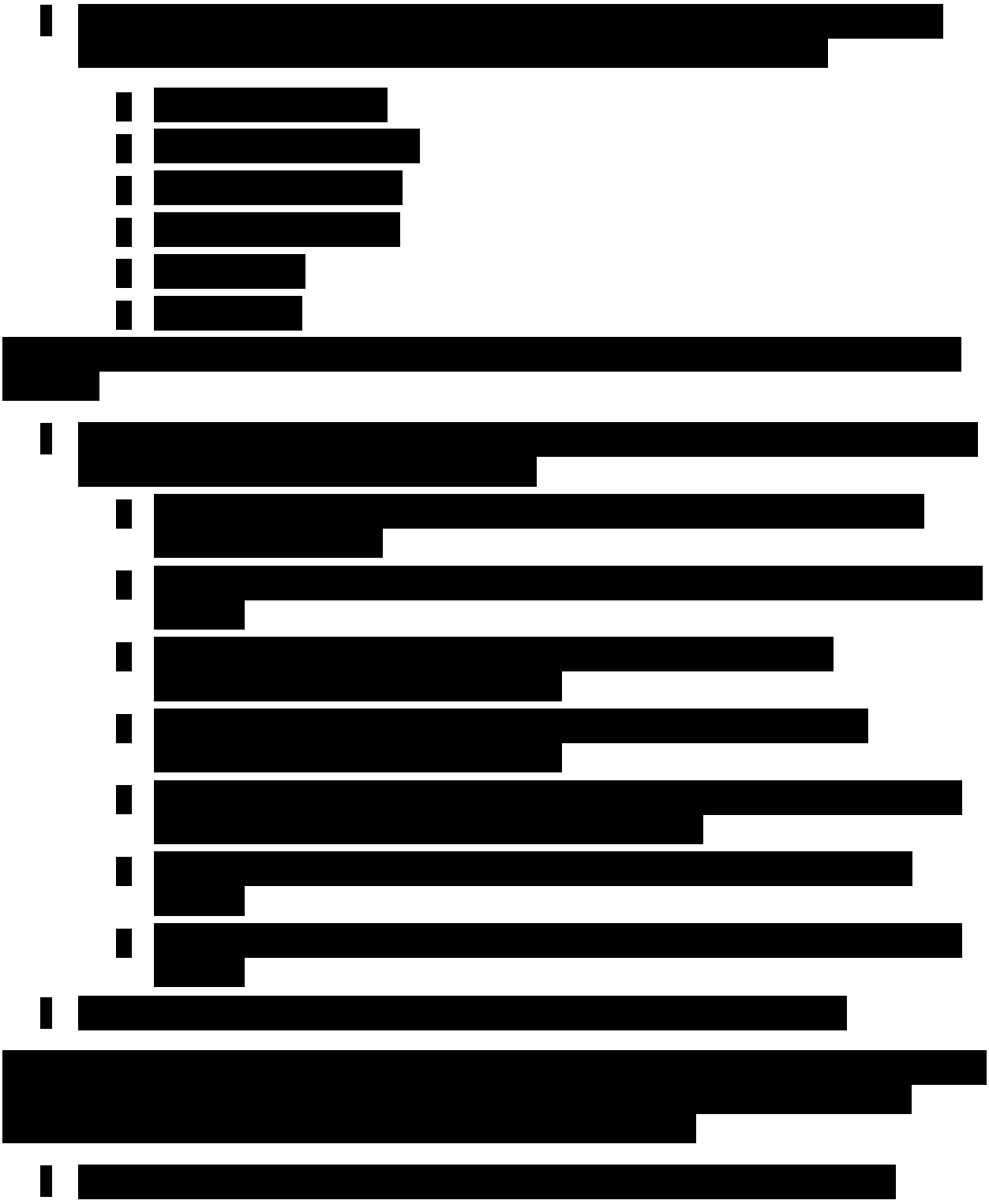
Table 4: Motor Milestone Development Survey (WHO-MGRS)

WHO-MGRS Item	**Performance Criteria**
Sitting without support	[redacted]
Hands-and-knees crawling	
Standing with assistance	
Walking with assistance	
Standing alone	
Walking alone	

[redacted]

[redacted]

6.4.3.2. Patients with bi-allelic SMN1 deletions and 3 copies of SMN2



[redacted]

[redacted]

6.5. VISIT WINDOWS

For efficacy analyses, the time windows specified in Table 5.1 (2-copy cohort) and Table 5.2 (3-copy cohort) describe how efficacy data will be assigned to protocol-specified time points for each cohort. For visits occurring prior to and including the Day 72 visit, data will be assigned to protocol-specified time points based on study day (as defined in Section 6.2.2). For visits occurring after the Day 72 visit, data will be assigned to protocol-specified time points based on patient's age at the visit/event (as defined in Section 6.2.8).

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 5.1: Analysis Time Windows of Bayley Scales, WHO-MGRS, CHOP-INTEND and CMAP – 2 copy cohort

Scheduled Visit	Nominal Days (Study Day [up to Month 2] or Age [after Month 2])	Acceptable Analysis Window (Min Day – Max Day)	Window applicable to:
Baseline	-	< 1	All assessments
Day 7	7	4 to 10	CHOP-INTEND only
Day 14	14	11 to 17	CHOP-INTEND only
Day 21	21	18 to 24	CHOP-INTEND only
Day 30/Month 1	30	27 to 33	All assessments

Scheduled Visit	Nominal Days (Study Day [up to Month 2] or Age [after Month 2])	Acceptable Analysis Window (Min Day – Max Day)	Window applicable to:
Month 2	60	46 to 74	All assessments
Age 3 Months	90	75 to 105	All assessments
Age 6 Months	180	152 to 208	All assessments
Age 9 Months	270	242 to 298	All assessments
Age 12 Months	360	332 to 388	All assessments
Age 15 Months	450	422 to 478	All assessments
Age 18 Months	540	540 to 569	All assessments

Table 5.2: Analysis Time Windows of Bayley Scales, WHO-MGRS, CHOP-INTEND and CMAP – 3 copy cohort

Scheduled Visit	Nominal Days (Study Day [up to Month 2] or Age [after Month 2])	Acceptable Analysis Window (Min Day – Max Day)	Window applicable to:
Baseline	-	< 1	All assessments
Day 7	7	4 to 10	CHOP-INTEND only
Day 14	14	11 to 17	CHOP-INTEND only
Day 21	21	18 to 24	CHOP-INTEND only
Day 30/Month 1	30	27 to 33	All assessments
Month 2	60	46 to 74	All assessments
Age 3 Months	90	75 to 105	All assessments
Age 6 Months	180	152 to 208	All assessments
Age 9 Months	270	242 to 298	All assessments
Age 12 Months	360	332 to 388	All assessments
Age 15 Months	450	422 to 478	All assessments
Age 18 Months	540	512 to 568	All assessments
Age 21 Months	630	602 to 658	All assessments
Age 24 Months	720	720 to 749	All assessments

The time windows specified in Tables 6.1 and 6.2 describe how safety data (laboratory, ECG, echocardiogram, vital signs) collected will be assigned to protocol-specified time points . For visits occurring prior to and including the Day 72 visit, data will be assigned to protocol-specified time points based on study day (as defined in Section 6.2.2). For visits following the Day 72 visit, data will be assigned to protocol-specified time points based on patient's age at the

visit (as defined in Section 6.2.8). Note that the windows for Day 37 and Day 51 are only applicable to patients enrolled at sites located in Japan.

If more than one record for a specific parameter is included in a time window, the record closer to the target day will be used. If there are two records equally distant to the target day, the latest one will be used in analyses.

Table 6.1: Analysis Time Windows of Safety Assessments – 2-copy cohort

Scheduled Visit	Nominal Days (Study Day [up to Day 72] or Age [after Day 72])	Acceptable Analysis Window (Min Day – Max Day)
Baseline	-	< 1
Day 2	2	1 to 3
Day 7	7	4 to 10
Day 14	14	11 to 17
Day 21	21	18 to 24
Day 30/Month 1	30	27 to 33
Day 37	37	34 to 40
Day 44	44	41 to 47
Day 51	51	48 to 54
Day 60/Month 2	60	57 to 63
Day 72	72	69 to 75
Age 3 Months	90	76 to 104
Age 6 Months	180	152 to 208
Age 9 Months	270	242 to 298
Age 12 Months	360	332 to 388
Age 15 Months	450	422 to 478
Age 18 Months	540	540 to 569

Table 6.2: Analysis Time Windows of Safety Assessments – 3-copy cohort

Scheduled Visit	Nominal Days (Study Day [up to Day 72] or Age [after Day 72])	Acceptable Analysis Window (Min Day – Max Day)
Baseline	-	< 1
Day 2	2	1 to 3
Day 7	7	4 to 10
Day 14	14	11 to 17
Day 21	21	18 to 24
Day 30/Month 1	30	27 to 33
Day 37	37	34 to 40
Day 44	44	41 to 47

Scheduled Visit	Nominal Days (Study Day [up to Day 72] or Age [after Day 72])	Acceptable Analysis Window (Min Day – Max Day)
Day 51	51	48 to 54
Day 60/Month 2	60	57 to 63
Day 72	72	69 to 75
Age 3 Months	90	76 to 104
Age 6 Months	180	152 to 208
Age 9 Months	270	242 to 298
Age 12 Months	360	332 to 388
Age 15 Months	450	422 to 478
Age 18 Months	540	512 to 568
Age 21 Months	630	602 to 658
Age 24 Months	720	720 to 749

The efficacy data of this study will be compared to the data from the PNCR network. The analysis visit of PNCR will be based on the Visit Month defined in the PNCR dataset.

Table 7: Analysis Time Windows for dataset from PNCR Network

Scheduled Visit	PNCR
Baseline	<= Enrollment Date
Day 30	Visit Month 1
Month 2	Visit Month 2
Month 3	Visit Month 3
Month 4	Visit Month 4
Month 6	Visit Month 6
Month 9	Visit Month 9
Month 12	Visit Month 12
Month 15	Visit Month 15
Month 18	Visit Month 18
Month 21	Visit Month 21
Month 24	Visit Month 24

If a PNCR patient is missing efficacy data at Baseline, then the earliest non-missing assessment collected post-Baseline will be re-mapped and considered their Baseline value, and subsequent visits occurring within 18/24 months after the re-mapped Baseline value will be included in the analysis.

For example, if a PNCR patient is missing their HFMSE assessment at Visit Month 0 and Visit Month 2 and their first non-missing HFMSE assessment occurred at Month 4, then their data would be re-mapped as follows:

- Visit Month 4 assessment is considered Baseline
- Month 6 is re-mapped as Month 2 (visit occurred 2 months after Baseline)
- Month 9 is remapped as Month 5 (visit occurred 5 months after Baseline)
- Month 12 is remapped as Month 8 (visit occurred 8 months after Baseline)
- Etc.

Safety data, such as laboratory results, vital signs, ECGs, echocardiograms, and physical exams will be assessed by date and study day. For change from baseline analyses the value associated with the scheduled visit will be used. For summaries of shifts from baseline and potentially significant values all values will be considered.

6.6. POOLING OF CENTERS

This is a multiple-center study. Given that the expected enrollment at any individual site is expected to be fewer than 3 patients, patients in each cohort from all sites will be pooled in the analyses for the primary, secondary, and exploratory efficacy endpoints. A site effect or a site-treatment interaction will be not examined because the sites selected for this study are all high-level treatment centers with extensive experience in the management of SMA and, as such, are expected to follow care management strategies closely aligned with the published standards of care.

6.7. SUBGROUPS

Patients with the *SMN2* gene modifier mutation (c.859G>C) will be evaluated separately as part of subgroup analyses. The *SMN2* copy number for these patients will determine how long they will be followed in the study (until 18 months of age for those with *SMN2*=2, until 24 months of age for those with *SMN2*=3), and the relevant set of primary, secondary, and exploratory endpoints that will be assessed (as detailed in Sections 6.4.1-6.4.3 of the SAP). The full set of efficacy endpoints (primary, secondary, exploratory) relevant for each *SMN2* copy number will be summarized for these subgroups of patients. All summaries will be limited to descriptive statistics.

Safety data for these subgroups of patients will be summarized by cohort together with the patients who do not have these mutations.

6.8. INDETERMINATE DATA

CMAP area values greater than 100 mVmsec will be considered to be indeterminate, and will be excluded from all relevant analyses. The indeterminate values will still be present in the data listings.

7. DEMOGRAPHIC, OTHER BASELINE CHARACTERISTICS

AND MEDICATION DATA

Demographics, other baseline characteristics, and medication data will be summarized separately for each cohort.

7.1. PATIENT DISPOSITION AND WITHDRAWALS

Patient disposition will be summarized as follows:

- The number of patients screened
- The number of patients who failed screening
- The number (%) of patients in the Enrolled Population
- The number (%) of patients in the Safety Population
- The number (%) of patients in the ITT Population
- The number (%) of patients in the EC Population
- 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. For the 3-copy cohort only, the following will also be summarized:
    - The number (%) of patients who completed the study within the 24 months of age window (≤749 days old)
    - The number (%) of patients who completed the study outside of the 24 months of age window (>749 days old)
- 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).

A listing of patient disposition will be provided.

The inclusion/exclusion criteria for the patients who failed screening and all protocol deviation data will be listed. A separate listing will be produced which summarizes all protocol deviations related to COVID-19.

7.2. DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

Demographics and baseline characteristics will be summarized for the ITT Population using descriptive statistics.

Demographics include age, gender, ethnicity, race, weight, and length.

- Age (days) = (dose date - date of birth +1)

- Length (cm) = Length (in inches) * 2.54
- Weight (kg) = Weight (in lbs) * 0.4536

Age will be summarized as a continuous variable as well as using the following groupings:

- Newborn (age 0-27 days)
- Infant and toddler (age 28 days – 23 months)

Race will be summarized as captured in the eCRF. Patients with multiple races will be summarized as 'Multiple'.

Baseline disease characteristics include the following:

- Age at SMA diagnosis for patients who were diagnosed after birth (i.e., date of birth is prior to date of SMA diagnosis), calculated as:

  (SMA diagnosis date - date of birth +1)
- Familial history of SMA including affected siblings or parent carriers
- Modality of SMA diagnosis
- Gestational age at birth
- Length/weight and head circumference at birth

All demographics and baseline characteristics will be listed.

7.3. MEDICAL HISTORY AND CONCOMITANT DISEASES

A summary table of the number and percentage of patients by medical history system organ class (SOC) and preferred term (PT) will be produced for the ITT Population. Medical history will be sorted alphabetically by SOC and PT using the MedDRA coding dictionary. For the summary tables, a patient will be counted at most once within a PT and/or SOC.

Listings for medical history, SMA medical history, and SMA diagnosis will also be provided.

7.4. MEDICATIONS

7.4.1. Prior and Concomitant Medication

All prior and concomitant medications will be classified using the ATC classification and preferred drug names from the WHO Drug Dictionary. Prior medications are defined as medications starting from 2 weeks prior to infusion of AVXS-101. Concomitant medications are defined as medications ongoing at time of the infusion of AVXS-101 or starting after the infusion through the last study visit.

Prior and concomitant medications will be summarized separately by ATC level 2 and preferred drug name for the ITT Population.

A listing for prior and concomitant medications will also be provided.

7.4.2. Specific Medication Subgroup

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

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 each day from Day 1 to Day 30, then on Day 31, dose is lowered to 0.5 mg/kg until Day 42. On Day 43, the dose is lowered to 0.25 mg/kg and continues until Day 56 when prednisolone dosing stops. For this patient,

Total Cumulative Dose= (1.0 mg/kg x 30 days) +
(0.5 mg/kg x 12 days) +
(0.25 mg/kg x 14 days) = 39.5 mg/kg*days

Exposure will be summarized using descriptive statistics for the Safety Population. A listing of these data will also be provided.

7.4.3. Measurement of Treatment Compliance

Refer to Section 9.2 of this SAP.

7.4.4. Other Therapies

Non-medication therapies and procedures will be defined as “Prior” and/or “Concomitant”. Prior non-medication therapies/procedures are defined as therapies or procedures started prior to infusion of AVXS-101. Concomitant non-medication therapies/procedures are defined as therapies or procedures ongoing at time of infusion of AVXS-101 or started after the infusion.

Non-medication therapies/procedures will not be summarized in tables. A listing of these data will be provided.

8. EFFICACY

An independent Central Reviewer will evaluate recorded videos of segments of assessments conducted at clinic visits that demonstrate developmental milestone achievement as determined by the clinical evaluator at the site. The reviewer will determine whether the video confirms evidence of the developmental milestone achievement as assessed by the clinical evaluator.

The analysis of all primary, secondary, and exploratory efficacy endpoints will be conducted using the ITT Population. Analyses may be repeated using the EC Population as sensitivity analyses in the event that the EC Population contains a different set of patients as compared to the ITT Population. If required by regulatory request, efficacy analyses may also be conducted using the population of patients who completed the study prior to COVID-19 pandemic disruptions arising.

Efficacy endpoints will be summarized separately for each cohort.

8.1. PRIMARY EFFICACY ENDPOINTS AND ANALYSIS

8.1.1. Patients with bi-allelic *SMN1* deletions and 2 copies of *SMN2*

The developmental milestone of sitting independently (as defined by BSIDv03 Gross Motor Subtest Item #26) will be assessed at every visit until attainment of the milestone. This developmental milestone will also be assessed at the 18 months of age visit, regardless of previous attainment, as the Bayley Scales do not necessarily require the child to repeat previously attained milestones.

The number and percentage of patients who demonstrate the developmental milestone of independent sitting for at least 30 seconds at any point up to and including the 18 months of age visit (up to 569 days of age) will be summarized. A one-sided exact binomial test will be used to test the null hypothesis of $p \leq 0.1\%$ at significance level of 0.025. Furthermore, the corresponding one-sided 97.5% confidence interval (CI) will be estimated by the exact method for binomial proportions.

Additionally, for patients who demonstrate the developmental milestone of functional independent sitting, the age at which they first demonstrated the developmental milestone will be summarized using descriptive statistics.

The number and percentage of patients who demonstrate the developmental milestone of functional independent sitting within the 99th percentile for normal development (≤279 days of age [4]), at an age older than the 99th percentile for normal development, and who do not demonstrate the developmental milestone at all will be presented.

If necessary, an additional set of sensitivity analyses may be conducted to include patients who demonstrated the developmental milestone of sitting independently after 18 months of age but within 3 months (ie, before 22 months of age; age 570 - 659 days) due to COVID-19 disruptions. These analyses will only be conducted in the event that COVID-19 disruptions impact the assessment of the primary endpoint in these patients.

In the first sensitivity analysis, the proportion of patients that achieved the independent sitting developmental milestone at any post-baseline visit will be summarized (considering the expanded 3-month window after Month 18 as defined above). The corresponding p-value from a one-sided exact binomial test will be computed for the comparison against PNCR. The difference

in the proportions between AVXS-101 and PNCR will be summarized and the exact 95% CI will be provided accordingly. The summary and statistical analysis will be conducted using the ITT Population.

In the second sensitivity analysis, the time from dosing to achieving the independent sitting developmental milestone will be summarized for patients in the ITT Population. The study day will be used as the time variable. Age at baseline will be used as a covariate. The probability of the ability to sit independently over time will be estimated using Kaplan-Meier methods. A Cox regression model will be used to obtain the hazard ratio with 95% CI for AVXS-101 to compare to the data from PNCR. In this analysis, patients who demonstrated the milestone within the expanded 3-month window after Month 18 will be censored at their last attended visit prior to 569 days of age.

8.1.2. Patients with bi-allelic SMN1 deletions and 3 copies of SMN2

The developmental milestone of standing alone for at least 3 seconds (as defined by BSIDv03 Gross Motor Subtest Item #40) will be assessed at every visit until attainment of the milestone. This milestone will also be assessed at the 24 months of age visit, regardless of previous attainment, as the Bayley Scales do not necessarily require the child to repeat previously attained milestones. Patients from the PNCR database will be considered to have achieved the endpoint of ability to stand alone without support for at least 3 seconds if they have a score of 2 on item #19 of the HFMSE. The HFMSE definition for "stands alone without support for at least 3 seconds" is the same as the definition that is used for BSIDv03 item #40.

The number and percentage of patients in the study and in the population-matched control cohort of the PNCR network who demonstrate the developmental milestone of standing alone for at least 3 seconds at any visit will be summarized. The actual proportions as well as the difference of the proportions between two data sources will be summarized and the exact 95% CI will be provided.

The corresponding p-value from a 2-sided Fisher's exact test with $\alpha = 0.05$ will be computed for the comparison between AVXS-101 and PNCR data.

For patients who demonstrate the developmental milestone of standing alone, the age at which they first demonstrated the developmental milestone will be summarized in both days and months using descriptive statistics. Additionally, the 95% CI for the median age of milestone achieved in both days and months will be summarized.

The number and percentage of patients who demonstrate the developmental milestone of standing alone within the 99th percentile for normal development (≤514 days of age [4]), at an age older than the 99th percentile for normal development, and who do not demonstrate the developmental milestone at all will be presented.

If necessary, an additional set of sensitivity analyses may be conducted to include AVXS-101 patients who demonstrated the developmental milestone of standing alone after but within 3 months (ie, before 28 months of age; age 750 - 839 days) due to COVID-19 disruptions. These analyses will only be conducted in the event that COVID-19 disruptions impact the assessment of the primary endpoint in these patients.

In the first sensitivity analysis, the proportion of patients that achieved the standing alone developmental milestone at any post-baseline visit will be summarized (considering the expanded 3-month window after Month 24 as defined above). The corresponding p-value from a 2-sided Fisher's exact test will be computed for the comparison against PNCR. The difference in the proportions between AVXS-101 and PNCR will be summarized and the exact 95% CI will be provided accordingly. The summary and statistical analysis will be conducted using the ITT Population.

In the second sensitivity analysis, the time from dosing to achieving the standing alone developmental milestone will be summarized for patients in the ITT Population. The study day will be used as the time variable. Age at baseline will be used as a covariate. The probability of the ability to stand alone over time will be estimated using Kaplan-Meier methods. A Cox regression model will be used to obtain the hazard ratio with 95% CI for AVXS-101 to compare to the data from PNCR. In this analysis, patients who demonstrated the developmental milestone within the expanded 3-month window after Month 24 will be censored at their last attended visit prior to age of 749 days.

8.2. SECONDARY EFFICACY ENDPOINTS AND ANALYSES

8.2.1. Patients with bi-allelic SMN1 deletions and 2 copies of SMN2

8.2.1.1. Proportion of patients surviving and not requiring permanent ventilation in the absence of acute illness or perioperatively assessed at 14 months of age

The number and percentage of patients in the study and in the population-matched control cohort of the PNCR network surviving event-free to 14 months of age will be summarized. Patients who terminate the study prior to reaching 14 months of age for any reason will be censored at the point of withdrawal. The actual proportions as well as the difference of the proportions between two data sources will be summarized and the exact 95% CI of the difference will be provided.

The corresponding p-value from a 2-sided Fisher's exact test with $\alpha = 0.05$ will be computed for the comparison between AVXS-101 and PNCR data.

A Kaplan-Meier survival curve will be produced.

If necessary, an additional sensitivity analysis may be conducted to consider patients who were lost to follow-up prior to reaching 14 months of age. In this analysis, these patients will not be considered treatment failures and instead will be censored at their last attended visit prior to reaching 14 months of age. This analysis will be based on the ITT Population. If no patients were

lost to follow-up prior to reaching 14 months of age, then this sensitivity analysis will not be conducted.

8.2.1.2. Proportion of patients achieving the ability to maintain weight at or above the third percentile without the need for non-oral/mechanical feeding support at any visit up to 18 months of age

The number and percentage of patients who maintain the ability to maintain weight at or above the third percentile without the need for non-oral/mechanical feeding support at any visit up to the 18 months of age visit will be summarized. A one-sided exact binomial test will be used to test the null hypothesis of $p \leq 0.1\%$ at significance level of 0.025. Furthermore, the corresponding one-sided 97.5% CI will be estimated by the exact method.

An additional analysis will be conducted which summarizes the number and percentage of patients who meet the ability to thrive endpoint at 18 months of age as defined in Section 6.4.2.1, along with the corresponding 97.5% binomial CI. The observed proportion will be compared to zero using a 1-sided exact binomial test [2, 5]. To make computation of the p-value possible, the value of 0.1% will be used in place of a literal zero.

If necessary, an additional sensitivity analysis may be conducted to include patients who demonstrated the ability to maintain weight at or above the third percentile without the need for non-oral/mechanical feeding support after 18 months of age but within 3 months (ie, before 22 months of age; age 570 - 659 days) due to COVID-19 disruptions. In this analysis, the proportion of patients that achieved ability to maintain weight at or above the third percentile without the need for non-oral/mechanical feeding support will be summarized (considering the expanded 3-month window after Month 18). The corresponding p-value from a one-sided exact binomial test will be computed for the comparison against PNCR. The difference in the proportions between AVXS-101 and PNCR will be summarized and the exact 97.5% CI will be provided accordingly. The summary and statistical analysis will be conducted using the ITT Population. These analyses will only be conducted in the event that COVID-19 disruptions impact the assessment of the secondary endpoint in these patients.

8.2.2. Patients with bi-allelic SMN1 deletions and 3 copies of SMN2

8.2.2.1. Proportion of patients achieving the ability to walk alone at any visit up to 24 months of age

The developmental milestone of walking alone with coordination (at least 5 steps; as defined by BSIDv03 Gross Motor Subtest Item #43) will be assessed at each visit until attainment of the milestone. This milestone will also be assessed at the 24 months of age visit, regardless of previous attainment, as the Bayley Scales do not necessarily require the child to repeat previously attained milestones.

The number and percentage of patients in the study and in the population-matched control cohort of the PNCR network who demonstrate the ability to walk alone at any point will be summarized. The actual proportions as well as the difference of the proportions between two data sources will be summarized and the exact 95% CI will be provided.

The corresponding p-value from a two-sided Fisher's exact test with $\alpha = 0.05$ will be computed for the comparison between AVXS-101 and PNCR.

For patients who demonstrate the developmental milestone of walking alone, the age at which they first demonstrated the developmental milestone will be summarized in both days and months using descriptive statistics. Additionally, the 95% CI for the median age of milestone achieved in both days and months will be summarized.

The number and percentage of patients who demonstrate the developmental milestone of walking alone within the 99th percentile for normal development (≤534 days of age [4]), at an age older than the 99th percentile for normal development, and who do not demonstrate the developmental milestone at all will be presented.

If necessary, an additional set of sensitivity analyses may be conducted to include AVXS-101 patients who demonstrated the milestone of walking alone after 24 months of age but within 3 months (ie, before 28 months of age; age 750 – 839 days) due to COVID-19 disruptions. These analyses will only be conducted in the event that COVID-19 disruptions impact the assessment of the secondary endpoint in these patients.

In the first sensitivity analysis, the proportion of patients that achieved the walking alone developmental milestone at any post-baseline visit will be summarized (considering the expanded 3-month window after Month 24). The corresponding p-value from a 2-sided Fisher's exact test will be computed for the comparison against PNCR. The difference in the proportions between AVXS-101 and PNCR will be summarized and the exact 95% CI will be provided accordingly. The summary and statistical analysis will be conducted using the ITT Population.

In the second sensitivity analysis, the time from dosing to achieving the walking alone developmental milestone will be summarized for patients in the ITT Population. The study day will be used as the time variable. Age at baseline will be used as a covariate. The probability of the ability to walk alone over time will be estimated using Kaplan-Meier methods. A Cox regression model will be used to obtain the hazard ratio with 95% CI for AVXS-101 to compare to the data from PNCR. In this analysis, patients who demonstrated the milestone within the expanded 3-month window after Month 24 will be censored at their last attended visit prior to when the Month 24 visit was scheduled to occur.

8.3. EXPLORATORY EFFICACY ENDPOINTS AND ANALYSES

8.3.1. Patients with bi-allelic SMN1 deletions and 2 copies of SMN2

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]entilatory support during the entire study will be summarized using descriptive statistics.

8.3.2. Patients with bi-allelic SMN1 deletions and 3 copies of SMN2

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

8.3.2.7. [redacted]

[redacted]

[redacted]

[redacted]

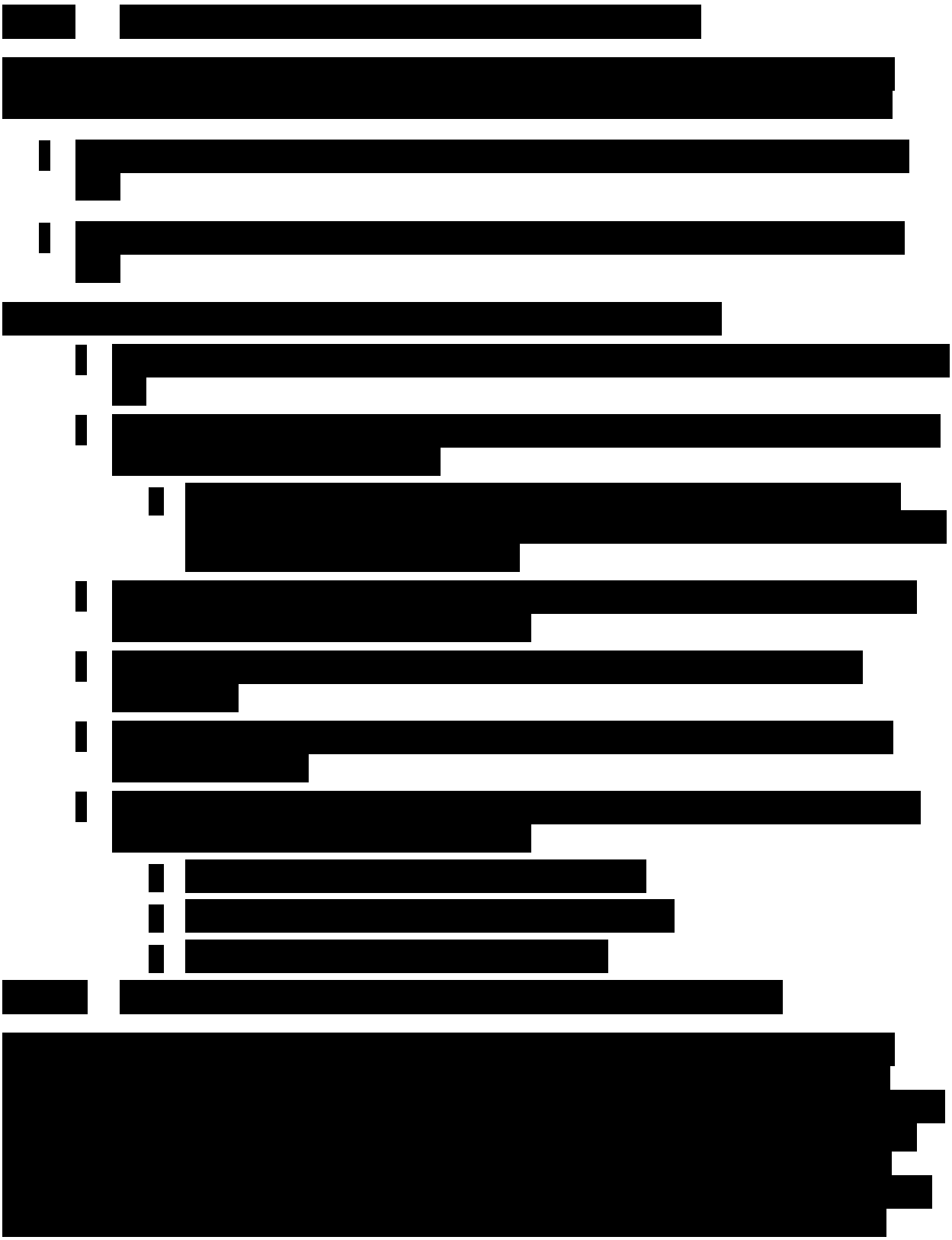
[redacted]

[redacted]

[redacted]

[redacted]

[redacted]



[redacted]

8.4. OTHER EFFICACY ENDPOINTS AND ANALYSES

Descriptive statistics for change from Baseline in CMAP amplitude, CMAP area , and surface temperature will be summarized by visit for each cohort using the ITT Population.

The change from Baseline for each CMAP parameter will be analyzed using a mixed model for repeated measurements (MMRM). The model will include the change from Baseline as the dependent variable, fixed effect of visit, and a covariate of Baseline value, age at Baseline, and interactions of Baseline value*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, and standard errors from model effects will be reported for each scheduled visit through Month 18 (2-copy cohort) and through Month 24 (3-copy cohort).

Additionally, the maximum change from Baseline in CMAP amplitude, CMAP area, and surface temperature will be summarized. The maximum change from Baseline for each CMAP parameter will be analyzed using an Analysis of Covariance model with maximum change from Baseline as the dependent variable and a covariate for the Baseline value.

9. SAFETY

The population used for safety analyses will be the Safety Population. Safety will be assessed on the basis of AEs, clinical laboratory data, physical examinations, non-invasive ventilatory support use, vital signs, cardiac assessments, and related exams. Safety data will not be compared with the PNCR natural history cohort as safety data were not collected in a comparable manner during the natural history study.

Safety data will be summarized separately for each cohort.

9.1. EXTENT OF EXPOSURE

The actual weight-based dose administered, total volume administered (planned and actual), total vg administered, and duration of infusion of the patients' infusion of study product on dosing date will be summarized. Duration of infusion is calculated as (end date/time of treatment administration – start date/time of treatment administration). Additionally, the number of patients with the entire volume infused and the number of patients with infusion interrupted will be summarized.

Total actual weight-based dose administered (vg/kg) will be calculated as follows:

- Actual volume administered (mL) * Lot Titer value (vg/mL) / Weight at Day 1 (kg)

Total vg administered will be calculated as follows:

- Actual volume administered (mL) * Lot Titer value (vg/mL)

A listing of each patient's infusion of AVXS-101 will be provided.

9.2. TREATMENT COMPLIANCE

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

- 100 * total volume administered / total planned volume administered

Treatment compliance will be summarized in a table and listed.

9.3. ADVERSE EVENTS / ADVERSE DRUG REACTIONS

9.3.1. Adverse Events / Adverse Drug Reactions

Adverse events will be coded using MedDRA. Treatment-emergent adverse events (TEAEs) are defined as any event that begins or worsens in severity on or after initiation of study product through the last study visit. If an incomplete or missing onset date was collected for an AE, the imputation method of missing data specified in Section 6.3 will be applied. An AE will be defined as a TEAE if its imputed onset date is on or after the date and time of infusion of study product through last study visit.

The CTCAE version 4.03 [3] will be used to report the severity of each AE. The classifications of CTCAE 4.03 are outlined in Table 8.

Table 8: Adverse Event Classification

Grade	Classification
1	Mild adverse event; did not require treatment
2	Moderate adverse event; resolved with treatment
3	Severe adverse event; inability to carry on normal activities; required professional medical attention
4	Life-threatening or permanently disabling adverse event
5	Fatal adverse event

The relationship to study treatment for each AE will be classified as 'Unrelated', 'Possibly Related', 'Probably Related', or 'Definitely Related'. AEs classified as 'Possibly', 'Probably' or 'Definitely' related will be analyzed as 'Related' in the AE summaries.

9.3.2. Tabulations of Treatment-Emergent Adverse Events

Adverse event data will be summarized and presented using primary MedDRA SOCs and PTs. The SOCs will be presented in alphabetical order and the PTs will be presented in alphabetical order within each SOC.

9.3.2.1. Adverse Event Overview

An overview of AEs will be presented consisting of the number and percentage of patients experiencing at least one event for the following AE categories:

- Any TEAE
- TEAEs related to study treatment
- TEAEs of Grade 3 severity or higher
- Serious TEAEs
- Serious and related TEAEs
- TEAEs leading to discontinuation of patient from study
- TEAEs leading to death

9.3.2.2. Adverse Events by SOC and PT

The following summaries of AEs will be generated:

- Incidence of TEAEs
- Incidence of TEAEs related to study treatment

- Incidence of Serious TEAEs
- Incidence of Serious TEAEs by relationship to study treatment
- Incidence of TEAEs with Grade 3 or Grade 4 severity
- Incidence of TEAEs leading to discontinuation from study
- Incidence of TEAEs leading to death
- Incidence of TEAEs leading to death by relationship
- Incidence of non-serious TEAEs

For all AE summaries, the number of TEAEs and the number and percentage of patients experiencing TEAEs will be tabulated according to SOC and PT. Patients reporting more than one AE for a given PT will be counted only once for that term, although each event will be counted individually. Patients reporting more than one AE within an SOC will be counted only once for that SOC.

Listings of TEAEs, related TEAEs, serious AEs (including pre-treatment events), serious AEs related to study treatment, AEs leading to study discontinuation, AEs leading to death, and TEAEs of special interest will be provided.

9.3.2.3. Adverse Events by PT

The number of TEAEs and the number and percentage of patients experiencing TEAEs will be summarized according to PT and sorted by overall frequency. Similar summaries will be provided for Grade 3 and Grade 4 TEAEs and TEAEs related to study drug/product.

9.3.2.4. Adverse Events of Special Interest

The following specific TEAEs of special interest, which are primarily defined by using Standardized MedDRA queries (SMQ), will be summarized:

- Hepatotoxicity, identified via the following SMQ:
  - Hepatic disorders (SMQ)
- Thrombocytopenia, identified via the following CMQ:
  - Transient thrombocytopenia (CMQ)
- Cardiac events, identified via the following SMQs:
  - Ischemic heart disease (SMQ)
  - Cardiomyopathy (SMQ)
  - Cardiac arrhythmias (SMQ)
  - Embolic and thrombotic events (SMQ)

- Myocardial infarction (SMQ)

- Thrombotic microangiopathy, identified via the following approach:
  - Criteria #1: cases with **any one** of the following PTs: thrombotic microangiopathy **OR** haemolytic uraemic syndrome **OR** atypical haemolytic uraemic syndrome **(i.e. A).**
  - Criteria #2: cases with **at least one PT** from **EACH** of the following SMQs representing thrombocytopenia, hemolysis and relevant renal events respectively:
    - Haematopoietic thrombocytopenia (SMQ) **(i.e. B)**
    - Haemolytic disorders (SMQ) **(i.e. C)**
    - Acute renal failure (SMQ)**(i.e. D) OR** Renovascular disorders (SMQ) **(i.e. E)**

- Sensory abnormalities suggestive of ganglionitis, identified via the following CMQ:
  - DRG Cell Inflammation (CMQ)

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. AEs of interest will be summarized by SOC and PT overall.

9.3.2.5. Adverse Events by Maximum Severity

TEAEs will be summarized by maximum CTCAE grade of each PT. Each PT will be assigned to a grade level, as assessed by the investigator, based on the CTCAE version 4.03 for grading severity of AEs. A patient who has an AE with unknown grade will be counted in the severity grade level category of "unknown". If the patient has another occurrence of the same event with a grade present, then the patient will be counted under the maximum grade that is present for the AE.

9.3.2.6. Adverse Events by Maximum Relationship

TEAEs will be summarized by maximum relationship of each PT to study product (AVXS-101), as assessed by the investigator. A patient who has an event with an assessment of 'Possibly Related', 'Probably Related' or 'Definitely Related', then the event will be summarized as 'Related' in the table. A patient who has an event with the assessment of 'Unrelated' will be summarized as 'Not Related' in the table. If a patient has more than one occurrence of the same event, and one is 'related' and the other is 'not related', then the 'related' event is considered to be the one having the maximum relationship to study drug. If a patient has an AE with unknown relationship, then the patient will be counted in the relationship category of "unknown". The only exception is if the patient has another occurrence of the same AE with a relationship present. In this case, the patient will be counted under the maximum relationship category.

9.4. LABORATORY EVALUATIONS

9.4.1. Variables and Criteria Defining Abnormality

Hematology variables: hematocrit, hemoglobin, red blood cells, white blood cells, platelets, neutrophils, lymphocytes, monocytes, basophils, and eosinophils.

Clinical Chemistry variables: gamma glutamyl transferase (GGT), AST, ALT, serum total bilirubin, direct bilirubin, albumin, glucose, total creatinine kinase, creatine kinase-MB (CK-MB), troponin I, creatinine, BUN, electrolytes (sodium, potassium, chloride, CO_2), and alkaline phosphatase.

Urinalysis variables: color, clarity/turbidity, pH, specific gravity, glucose, ketones, nitrites, leukocyte esterase, bilirubin, blood, protein, red blood cells, white blood cells, squamous epithelial cells, casts, crystals, bacteria, and yeast.

Viral testing variables: human immunodeficiency virus, hepatitis B, hepatitis C, and Zika virus. These samples will be collected from the patient's biological mother at Screening; Zika testing is only required for mothers with clinical suspicion of Zika virus.

Immunology variables: serum antibody to AAV9 (anti-AAV9) and serum antibody to SMN (anti-SMN). This sample will be collected from either the patient or the patient's biological mother at Screening.

Diagnostic re-confirmation testing variables: *SMN1* deletions/mutations, *SMN2* copy number, absence of c.859G>C (technical result), variant type.

Saliva, urine, and stool collection (patients enrolled at Japan sites only): saliva, urine, and stool samples will be collected at Day 2, Day 3, Day 7, Day 14, Day 21, and Day 30. Additionally, urine and stool samples will be collected over a 24- hour period immediately post-infusion, and for 24-hour periods at Day 30 and Day 60, in accord with the laboratory manual for viral shedding studies.

The Criteria for Potentially Clinically Significant (PCS) values are generally based upon CTCAE version 4.03 [3] criteria for Grade 2 or higher AEs. Age-related differences in normal ranges of some analytes in pediatric (vs adult) populations have been taken into account in defining PCS values.

The Criteria for PCS laboratory findings are maintained outside of this SAP.

9.4.2. Statistical Methods

Clinical laboratory tests analyzed by the central laboratory will be summarized at the scheduled visits in accordance with the Schedule of Study Assessments. Laboratory tests analyzed by the local laboratory will not be included in any tables but will be provided in listings.

Values at each scheduled visit as well as change from baseline values to each post-baseline visit will be summarized for each protocol-specified chemistry and hematology laboratory parameter using descriptive statistics. Urinalysis parameters viral testing parameters, immunology variables, and diagnostic re-confirmation testing variables will be summarized in listings only.

Chemistry and hematology 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 study period for each cohort 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 meeting the specified criteria for PCS laboratory values will be summarized. A post-baseline value must be more extreme than the baseline value to be considered a PCS finding. Listings will be provided to present all lab values for the patients meeting PCS criteria during treatment.

For hemoglobin and the liver function enzyme (LFE) tests of ALT, AST, alkaline phosphatase, and GGT, the number and percentage of patients with a maximum CTCAE Grade of 1, 2, 3, or 4 (for ALT or AST) or Grade 3 or 4 (for alkaline phosphatase and GGT), (as defined by the central laboratory and based on CTCAE v4.03 [3]) at any post-baseline visit (regardless of the baseline value) through the end of study will be summarized. For ALT and AST, the table 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 a Grade 3 laboratory abnormality. Accompanying listings of all ALT, AST, GGT, and alkaline phosphatase will be created for each cohort for any patients who had at least a Grade 1 ALT or AST value or at least a Grade 3 alkaline phosphatase of GGT value. A listing of hematology results will be provided for patients with hemoglobin abnormalities (defined as any Grade 3 or higher hemoglobin value).

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

- ALT >3x ULN, ALT >5x ULN, ALT >10x ULN, ALT >20x ULN
- AST >3x ULN, AST >5x ULN, AST >10x ULN, AST >20x ULN
- ALT or AST >3x ULN, ALT or AST >5x ULN, ALT or AST >8x ULN, ALT or AST >10x ULN, ALT or AST >20x ULN
- Total bilirubin >2x ULN, Total bilirubin >3x ULN
- ALT or AST >3x ULN and Total bilirubin >2x ULN
- ALT or AST >3x ULN and Total bilirubin >2x ULN and Alkaline phosphatase ≥2x ULN
- ALT or AST >3x ULN and Total bilirubin >2x ULN and Alkaline phosphatase <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 the last three combination categories, the values do not need to have been collected at the same assessment. 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.

9.5. SWALLOWING TEST

A swallowing test will be performed at screening, every 6 months starting at 6 months of age, and at EOS to determine if the patient has signs of aspiration. If the test is positive for aspiration, the parent or guardian of the patient will be recommended to use an alternative method to oral feeding. Once implemented, 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 listed. Actual use of non-oral feeding support will be quantified through the recording of volume, frequency of use, duration, and calories; these data will be listed.

9.6. VITAL SIGNS, WEIGHT AND LENGTH

Vital signs will include systolic blood pressure, diastolic blood pressure, heart rate, respiratory rate, pulse, temperature, and pulse oximetry. Body weight and length will also be assessed.

For each scheduled visit and time point, observed values as well as change from baseline values will be summarized using descriptive statistics.

In addition, vital sign results will be flagged as PCS if they meet the pre-specified criteria which are defined outside of this SAP. The number and percentage of patients meeting each PCS criterion at any post-Baseline visit will be summarized. A listing of all PCS vital sign values will be provided.

9.7. CARDIAC MONITORING

9.7.1. ECG

A 12-lead ECG will be conducted at the scheduled visits in accordance with the Schedule of Study Assessments. ECGs will be interpreted locally by a cardiologist for immediate safety evaluation. The ECG tracings or ECG machine data will also be collected for centralized review and interpretation by a cardiologist.

Observed values as well as change from baseline values will be summarized at each scheduled visit using descriptive statistics for HR, RR, PR, QRS, QT, and QTcF as measured by the central reviewer. Additionally, a summary of ECG interpretation (normal, abnormal - clinically insignificant, abnormal – potentially clinically significant) will be provided.

In addition, ECG results will be flagged as PCS if they meet the pre-specified criteria which are

defined outside of this SAP. The number and percentage of patients meeting each PCS criterion will be summarized. A listing of all PCS ECG values will be provided.

9.7.2. Echocardiogram

Echocardiograms will be conducted at Screening, at the 3, 6, 9, and 12 months of age visits, and every 6 months thereafter. Echocardiograms will be interpreted locally and results from the local interpretation (abnormal/normal, etc.) will be captured in the eCRF. Additionally, echocardiogram data will be provided to an external cardiologist for centralized review; this will be considered the primary echocardiogram source data.

Clinically significant, treatment-emergent findings should be reported as AEs. A listing of echocardiogram results including the ejection fraction (ejec fraction, %), the shortening fraction (frac short, %), findings (of aortic valve, mitral valve, tricuspid valve, and pulmonic valve), and summary report based on the centralized review will be provided for all screening and post-baseline visits for each cohort.

9.7.3. Holter Monitoring

Patients will have a 12-lead continuous Holter monitor attached at each visit as indicated in the schedule of assessments and the Holter will remain in place through 24 hours. Serial ECG data from the Holter monitor will be pulled in triplicate at time points of ‘0 hour’, ‘2 hour’, ‘4 hour’, ‘6 hour’, ‘8 hour’, ‘12 hour’, ‘24 hour’ and assessed by a central reviewer. The triplicate of each parameter for HR, RR, PR, QRS, QT, QTcF, as measured by the central reviewer, will be averaged for summaries at each time point.

The central reviewer will identify abnormal ECGs that are PCS, based on central reviewing guidelines provided by the vendor.

Holter data will be presented in a listing.

9.8. PHYSICAL EXAMINATION

Physical examinations will be conducted by the Investigator or Sub-Investigator at each scheduled visit, except Day -1. Treatment-emergent abnormal findings on physical exams will be recorded as AEs. Any post-infusion abnormal physical exam findings will be listed by patient with the corresponding result on the baseline physical exam.

9.9. PULMONARY EXAM

Patients will be assessed by a pulmonologist at the time points specified in the Schedule of Study Assessments and may be fitted with a non-invasive positive pressure ventilator (e.g., 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).

Pulmonary exam results will be presented in a listing.

9.10. VENTILATORY SUPPORT

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

10. HEALTH ECONOMICS

Hospitalization information from time of birth including number, duration of hospitalization, and reason for hospitalizations including primary ICD-10 codes if available will be listed.

11. INTERIM SAFETY REVIEWS

The DSMB/DMC is an independent multidisciplinary group consisting of clinicians and a biostatistician that, collectively, have experience in the management of patients with SMA and other diseases, and in the conduct and monitoring of randomized clinical studies with interim analyses. The DSMB/DMC will be chartered to oversee the safety of patients during the conduct of the study and will act in an advisory capacity to AveXis. A detailed description of the DSMB/DMC, its role in this study, and the timing of the scheduled reviews will be described in a DSMB/DMC Charter.

The DSMB/DMC will routinely convene on a quarterly basis to review emerging safety data from the study. Following each meeting, the DSMB/DMC will make a recommendation as to whether or not the accumulated safety data warrants a suspension or discontinuation of the study, a modification to the study, or any additional comments or recommendations related to safety. The DSMB/DMC will prepare and provide minutes of their meetings to AveXis who will provide copies to the regulatory authorities as appropriate.

12. CHANGES FROM ANALYSIS PLANNED IN PROTOCOL

The following additional endpoints have been added for both cohorts:

- Proportion of patients achieving the ability to remain independent of ventilatory support at 18 months of age (2-copy cohort) and at 24 months of age (3-copy cohort)

- Proportion of patients achieving the ability to remain independent of ventilatory support for the entire study duration

The following additional endpoints have been added for the 3-copy cohort:

- Summarization of highest milestone achieved (considering both WHO-MGRS and Bayley milestones).

- Summarization of the number and percentage of patients achieving a scaled score on the BSIDv03 Gross and Fine Motor Subtests ≥4 at each visit, at all visits combined, and at any visit up to and including the 24 months of age visit.

13. PROGRAMMING CONSIDERATIONS

13.1. DATA HANDLING AND TRANSFER

All case report form data will be entered into the [redacted] (CRO) validated database and an electronic audit trail of edits maintained. Information on laboratory data and ventilatory support may be imported to the database electronically. Further details regarding Data Handling and Transfer specifications are described in the [redacted] Data Management Plan.

13.2. DATA SCREENING

Beyond the data screening built into the [redacted] Data Management Plan, the CRO programming of analysis datasets, tables, figures, and listings (TFL) provides additional data screening. Presumed data issues will be output into SAS logs identified by the word "Problem" and extracted from the logs by a SAS macro and sent to Data Management.

Review of a pre-freeze TFL run on clean subjects and a post-freeze TFL run on the frozen database allow for further data screening prior to lock. The post-freeze TFL will be discussed with the sponsor in a data review meeting to identify any final data issues and seek corrections prior to database lock The CRO statistician and the sponsor must approve database lock.

13.3. VALIDATION

The CRO's quality control procedures will be documented separately in the study specific quality control plan.

14. REFERENCE LIST

1. Kolb SJ, Coffey CS, Yankey JW, Krosschell K, Arnold WD, Rutkove SB, et al. Natural history of infantile-onset spinal muscular atrophy. Ann Neurol. 2017;82(6):883-91.
2. Finkel RS, McDermott MP, Kaufmann P, Darras BT, Chung WK, Sproule DM, et al. Observational study of spinal muscular atrophy type I and implications for clinical trials. Neurology. 2014;83(9):810-7.
3. Cancer Therapy Evaluation Program, Common Terminology Criteria for Adverse Events, Version 4.03, DCTD, NCI, NIH, DHHS (http://ctep.cancer.gov), Publish Date: Jun 14, 2010.
4. WHO Multicentre Growth Reference Study Group. WHO Motor Development Study: Windows of Achievement for Six Gross Motor Development Milestones. Acta Paediatrica, 2006; Suppl 450: 86-95.

5. Sproule DM, Hasnain R, Koenigsberger D, Montgomery M, DeVivo DC, Kaufmann P. Age at onset predicts likelihood and rapidity of growth failure among infants and young children with spinal muscular atrophy types 1 and 2. Child Neurol. 2012; 27(7):845-851.
6. World Health Organization Child Growth Standards Technical Report (http://www.who.int/childgrowth/standards/Technical_report.pdf) Tables 38, 39, 49, and 50.
7. Mercuri E, Lucibello S, Perulli M, et al. Longitudinal natural history of type I spinal muscular atrophy: a critical review. Orphanet J Rare Dis 2020; 15: 84.
8. [Bayley N (2006)] Bayley scales of infant and toddler development, Third Edition (Bayley-III). San Antonio, TX: The Psychological Corporation.

15. AMENDMENT(S) TO THE STATISTICAL ANALYSIS PLAN

15.1. AMENDMENT 1 – SUMMARY OF CHANGES

All text referencing patients with bi-allelic SMN1 deletions and 4 copies of SMN2 was removed based on the protocol amendment.

General edits were made throughout the entire document in order to clarify wording.

The number of patients planned to be enrolled into the 2 copy cohort was updated from 15 to 14 throughout.

Section 1.1 was updated to clarify the derivation of Study Day.

Section 2.1 was updated to refer to 'the CRO selected by AveXis' rather than to [redacted].

Section 3.3.1.1 was updated to reflect additional exploratory efficacy endpoints defined per protocol.

Section 3.3.1.2 was updated to reflect additional exploratory efficacy endpoints defined per protocol. Section 3.4.1 was updated to detail the PNCR comparator assumptions for the secondary endpoints.

Section 3.10 was updated to only contain the study flowchart. The schedule of study assessments can be found in the study protocol.

Section 4.1 was updated to add additional details regarding the review of milestone videos.

Section 4.3.1 was updated to reflect additional exploratory efficacy endpoints defined per protocol.

Section 4.3.2 was updated to reflect additional exploratory efficacy endpoints defined per protocol.

Section 4.4.3 was updated to reflect the collection of head circumference.

Section 4.4.7 was updated to reflect the updated time points at which ECGs will be assessed during the study. QR and QTcB were removed and RR and QT were added to the list of ECG assessments being conducted.

Section 4.4.8 was updated to reflect the updated time points at which Echocardiograms will be assessed during the study.

Section 4.1.12 was added to reflect the time points at which twenty-four hour Holter Monitoring will be conducted during the study.

Section 5.2 was updated to describe how patients with 4 copies of *SMN2* would be handled for the ITT Population.

Section 5.4 was updated to document that the EOS visit can be conducted either remotely or at the study site.

Section 5.6 was updated to include a population of all screened patients, in order to summarize reasons for screen failure patients.

Section 6.1 was updated to state that efficacy summaries would only be produced for the EC Population if the EC Population consists of a subset of patients in the ITT Population.

Section 6.2.2 was updated to explain the derivation of study day.

Section 6.2.7 was added to describe the determination of *SMN2* copy number.

Section 6.2.8 was added.

Section 6.3 was updated to remove the imputation rules for incomplete stop dates (AEs and medications) and to remove text regarding “LOCF mixed components” imputation.

Section 6.4.1 was updated to remove the reference to motor milestones represented by Bayley items 19 and 22 as these milestones are not being video confirmed in this study; to remove the statement “Although the raw and scaled score of each scale/subtest will be captured in the eCRF, all scores used in the analysis will be derived programmatically”.

Section 6.4.2.1 was updated to include a sensitivity analysis of the ‘ability to thrive’ endpoint which includes swallow test results; to remove the swallow test component from the primary definition/analysis of the ability to thrive endpoint; to document the source of whether a patient has an acute reversible illness excluding perioperative ventilation; to clarify the derivation of

'ability to thrive'; to detail the 'worse case' approach to determining swallow test results for cases where there are multiple results available for a consistency tested at the same time point.

Section 6.4.2.2 was updated to clarify the definition of maintenance of achieved milestones.

Section 6.4.3.1 was updated to clarify the definition of survival without tracheostomy and to describe the scoring of CHOP-INTEND.

Section 6.4.3.2 was updated to clarify the definition of maintenance of achieved milestones.

Section 6.5 was updated to provide efficacy visit windows tailored to the protocol-defined windows, to provide visit windows for all safety assessments, to provide instruction on visit windowing for the PNCR data.

Section 7.1 was updated to include a summary of patients who discontinued due to AE by type of AE (non-fatal, serious fatal, non-serious fatal); to include a listing of screen failures; to include a listing of protocol deviations related to COVID-19; to clarify the definition of a completer patient.

Section 7.2 was updated to include a categorized summary of age and modality of SMA diagnosis, to summarize baseline demographics using the ITT population, to include a summary of age at SMA diagnosis.

Section 7.3 was updated to include a listing of SMA diagnosis data.

Section 7.4.1 was updated to clarify that ATC level 2 would be used to summarize prior and concomitant medications, to summarize medications using the ITT population.

Section 8 was updated to include the following text: If required by regulatory request, efficacy analyses may also be conducted using the population of patients who completed the study prior to COVID-19 pandemic disruptions arising.

Sections 8.1.1, 8.1.2, and 8.2.2.1 were updated to remove the summary of duration of the achieved milestone. The text related to date when milestone was lost has been removed as this information is not being captured in the clinical database. Sensitivity analyses which may be conducted to address COVID-19 disruptions were added. The forest plot and the summary of sitting milestones achieved for patients who do not demonstrate the milestone of independent sitting for at least 30 seconds was removed. A summary of the number and percentage of patients who demonstrate the milestone of functional independent sitting within the 99th percentile for normal development (9.2 months of age) was added.

Section 8.2 was updated to include sensitivity analyses to address COVID-19 disruptions.

Section 8.2.1.2 was updated to clarify the definition of the secondary endpoint of ability to thrive, and to include a sensitivity analysis of ability to thrive which includes swallow test results.

Section 8.3.1.1 was updated to summarize 97.5% CIs rather than 95% Cis; to include a summary of the minimum, median, and maximum age at earliest achievement for patients demonstrating each milestone. For each milestone, a summary of the percentage of patients who demonstrate the milestone within the 99th percentile (upper bound of WHO-MGRS window for normal development), who demonstrate the milestone at an age above the 99th percentile, and who do not demonstrate the milestone at all was added.

Sections 8.3.1.5 and 8.3.1.6 were updated to remove the statistical testing. The PNCR database does not contain data for the Bayley Scales of Infant and Toddler Development and so there is no way to compare to this data. Descriptive summaries of observed and change from baseline in Bayley Gross and Fine Motor scores were added.

Sections 8.3.1.6 and 8.3.2.6 were updated to clarify that the scaled score should be >=5.5 rather than >=7 for the analysis. The analysis approach was updated in order to summarize patients who achieve the scaled score at each individual visit, at all visits combined, and additionally at any visit.

Section 8.3.1.7 was updated to include spaghetti plots of Bayley scores over time.

Sections 8.3.1.8, 8.3.1.9, and 8.3.1.10 were updated to remove the condition that the analysis would be based on the subset of patients with a baseline CHOP-INTEND score <40, <50, and <58.

Sections 8.3.1.10 and 8.3.2.8 were updated to amend the summary of maintained milestones.

Section 8.3.1.10 was updated to include data handling rules for patients who achieved 3 consecutive CHOP INTEND scores >=58 and had subsequent CHOP INTEND assessments conducted; to reference the CHOP INTEND spaghetti plot.

Section 8.3.1.12 was updated to summarize the number and percentage of patients who were able to be assessed for the milestone at the 18 months of age visit. Patients who have the following reasons recorded for not maintaining the milestone will be considered to not have been able to be assessed: ‘patient not assessed’, ‘patient non-compliant’. Additionally, the number and percentage of patients who maintained the milestone per the CRF will be summarized.

Section 8.3.1.13 was updated to include a summary of the minimum, median, and maximum age at earliest achievement for patients demonstrating each milestone. For the milestones of sitting without support, standing alone, and walking alone, a summary of the percentage of patients who demonstrate the milestone within the 99th percentile (upper bound of WHO-MGRS window for normal development), who demonstrate the milestone at an age above the 99th percentile, and

who do not demonstrate the milestone at all was added. A listing of the BSIDv03 milestones which were demonstrated at Screening was added.

Section 8.3.1.14 was added to describe the analysis for the endpoint of 'proportion of patients achieving ability to remain independent of ventilatory support at 18 months of age'.

Section 8.3.2.1 was updated to summarize 97.5% CIs rather than 95% CIs.

Section 8.3.2.7 was updated to include spaghetti plots of Bayley scores over time.

Section 8.3.2.8 was updated to include statistical testing.

Section 8.3.2.11 was added to describe the analysis for the endpoint of 'proportion of patients achieving ability to remain independent of ventilatory support at 24 months of age'.

Section 8.4 was updated to include the summary and analysis of CMAP parameters.

Section 9.1 was updated to include the algorithm to calculate actual weight-based dose administered, total vg administered, and to provide further details on the data to be summarized in summary tables.

Section 9.2 was updated to clarify the compliance calculation is based on planned dose administered.

Section 9.3.2.2 was updated to include tables for non-serious TEAEs, serious TEAEs by relationship, and fatal TEAEs by relationship, to include listings of related TEAEs, serious AEs (including pre-treatment events), serious AEs related to study treatment, AEs leading to study discontinuation, AEs leading to death, and TEAEs of special interest.

Section 9.3.2.4 was updated to refer to the current list of AEs of special interest for the AVXS-101 program.

Section 9.3.2.6 updated terminology from 'Unrelated' to 'Not Related'.

Section 9.4.1 was updated to reflect the inclusion of troponin I, and to remove the tables defining PCS lab values as these are now maintained in a separate document and outside of the SAP.

Section 9.4.2 was updated to state a listing would be produced to summarize urine and stool sample data for patients enrolled at Japanese sites; also to update that hematology tests for patients with at least one Grade 3 hemoglobin value would be listed; to remove the in-text table of CTCAE grades (this will be defined by the central lab); to add elevated LFT criteria; to remove the summary of the number and percentage of patients with a total bilirubin value in the categories of ≤ULN, >ULN – <2 × ULN, and ≥2 × ULN at the last visit of the study for patients meeting the criterion of ALT <3 × ULN and total bilirubin ≥2 × ULN; to remove the summary of the ratio of indirect bilirubin to total bilirubin for patients meeting the criterion of ALT <3 ×

ULN and total bilirubin ≥2 × ULN; to remove the summary of the ratio at baseline and the ratio associated with the peak total bilirubin value during the study; to remove the summary of the number and percentage of patients with a ratio <0.75 and <0.50 for baseline and peak total bilirubin; to remove the listing of TEAEs (defined as PTs within the "Cholestasis and jaundice of hepatic origin" [broad search] SMQ, excluding PTs within the "Investigations" SOC) for patients with a Grade 3 or higher total bilirubin elevation, to include GGT in the summary of LFTs; to clarify the summarization of LFEs by CTCAE grade.

Section 9.4.3 was deleted.

Section 9.6 was updated to summarize vital sign PCS values occurring at any Post-Baseline visit; to remove the in-text table of PCS lab values as these are now maintained in a separate document and outside of the SAP.

Section 9.7.1 was updated to remove the in-text table of PCS ECG values as these are now maintained in a separate document and outside of the SAP; to update the list of ECG parameters being assessed in this study; to clarify that PCS ECG values would not be summarized by visit, to add, a summary of ECG interpretation (normal, abnormal - clinically insignificant, abnormal – potentially clinically significant).

Section 9.7.2 was updated to include details for the echocardiogram listing.

Section 9.7.3 was updated to clarify that Holter data would be listed.

Section 9.8 was updated to reflect the updated time points at which Echocardiograms will be assessed during the study.

Section 9.11 (CMAP) was moved to section 8 as this is an efficacy assessment.

Section 9.12 was added to reflect the summarization of Holter Monitoring data.

Section 11 was updated to reflect the current protocol wording regarding the DSMB/DMC.

Section 12 was updated to include the updated analysis approach for the ability to thrive endpoint for the 2-copy cohort; to include a new endpoint which assesses the proportion of patients achieving the ability to remain independent of ventilatory support at 18 months of age (2-copy cohort) and at 24 months of age (3-copy cohort)

Sections 13.1, 13.2, and 13.3 were updated to not refer to [redacted] for conducting the statistics and programming for this study.

Section 14 was updated to include references to the WHO Motor Development Study, the NeuroNext natural history study, and the PNCR natural history study,

In-text Table 1 was updated to reflect the current schedule of assessments per protocol v3.0.

In-text Tables 10 and 11 were updated to remove the normal range column and to reflect values using SI units. Additional very low and very high PCS ranges were added to cover the full spectrum of expected patient ages in the study.

In-text Table 13 was updated to correct the reference for potentially clinically significant values for body weight.

15.2. AMENDMENT 2 – SUMMARY OF CHANGES

Wording was clarified throughout regarding the secondary endpoint for the 2-copy cohort (ability to maintain weight at or above the third percentile without need for non-oral/mechanical feeding support at any visit up to 18 months of age).

Wording was clarified throughout for the primary endpoint for the 3-copy cohort ("stand alone" rather than "stand without support" as well as the secondary endpoint for the 3-copy cohort ("walk alone with coordination" instead of "walk alone").

Section 3.4.2 was updated to clarify how the subset of PNCR patients for statistical comparison to the 3-copy cohort was performed. Text was added to document that the achievement of a score of 2 on HFMSE items #19 and #20 was not restricted to occurrences up to 24 months of age in the natural history cohort (achievement of this score at any time while in the PNCR study was considered).

Section 3.11 was added to describe the 4-copy cohort.

Section 5.3 was updated to align the definition of the ITT population with the study protocol.

Section 6.1 was updated to align the definition of the ITT population with the study protocol.

Section 6.2.7 was updated to include the following text: If the genetic test results are indeterminate, then the patient will be considered to have the higher copy number indicated in the result. For example, a genetic result of "Atypical (between 2-3 copies)" would result in a patient being assigned to the 3 copy cohort.

Section 6.3 was updated to state that any CHOP-INTEND assessments which have been flagged as "invalid" on the eCRF will not be included in any analyses.

Section 6.7 was updated to align the definition of the ITT population with the study protocol.

Section 6.8 (Indeterminate data) was added.

Section 7.1 was updated to include a summary of the number (%) of patients who completed the study within the 24 months of age window (≤749 days old) and outside of the 24 months of age window (>749 days old).

Section 8.1.2 was updated to include a summary of the 95% CI for the median age of milestone achieved in both days and months.

Section 8.2.1.1 was updated to accurately reflect the final summary/analysis for this endpoint as part of the 2-copy cohort final run.

Section 8.2.2.1 was updated to include a summary of the 95% CI for the median age of milestone achieved in both days and months.

Section 8.3.2.1 was updated to include a summary of the age of each milestone achieved, and the 95% CI for the median age of each milestone achieved in both days and months; to include a table which summarizes the number and percentage of patients by highest milestone achieved.

Section 8.3.2.5 was updated to clarify that percentages will be based on the number of patients with available baseline BSIDv03 Gross or Fine Motor Subtest scores.

Section 8.3.2.6 was updated to include a summary of the number and percentage of patients achieving a scaled score on the BSIDv03 Gross and Fine Motor Subtests ≥4 at each visit, at all visits combined, and at any visit up to and including the 24 months of age visit.

Section 8.3.2.7 was updated to clarify that two sets of spaghetti plots of Bayley data will be produced: one set will include the Bayley normal range (+/- 2 SD) as a reference over time and the other set will not include the Bayley normal range; to add a listing summarizing the reason(s) why Bayley Scale assessments were not performed.

Section 8.3.2.8 was updated to include the production of a spaghetti plot of body weight over time, including the WHO third percentile values. An additional summary table was added which summarizes this same endpoint, however is based on any/all available weights recorded for the patient, and does not require weights to be recorded through the 24 months of age visit.

Section 8.3.2.10 was updated to include a summary of the 95% CI for the median age of each milestone achieved in both days and months; to include a table which summarizes the number and percentage of patients by highest milestone achieved.

Section 8.4 was updated to clarify the CMAP analyses are for both the 2-copy and 3-copy patient cohorts.

Section 9.3.2.4 was updated to reflect the current approach to identifying AESIs for the study.