COVER PAGE

Official Title:	Escalating Dose and Randomized, Controlled Study of Nusinersen (BIIB058) in Participants With Spinal Muscular Atrophy
NCT Number:	NCT04089566
Document Date:	Part A: 22 February 2021 Part B Infantile Onset: 19 July 2024 Part B Later Onset: 10 April 2024 Part C: 06 October 2023

DocuSign Envelope ID: ECFD44FD-BFB1-4502-A562-95FFA3B107E2

PART A

STATISTICAL ANALYSIS PLAN

Version No.: V1.0

Date: 22 February 2021

Author: [redacted]

Study Title: Escalating Dose and Randomized, Controlled Study of Nusinersen (BIIB058) in Participants With Spinal Muscular Atrophy

Name of Study Treatment: Nusinersen

Protocol No.: 232SM203

Study Phase 2/3

Confidential Information

The information contained in this Statistical Analysis Plan (SAP) is confidential and may not be reproduced or otherwise disseminated without the approval of Biogen, Inc.

DocuSign Envelope ID: ECFD44FD-BFB1-4502-A562-95FFA3B107E2

APPROVAL

This document has been reviewed and approved by:		
[redacted]	DocuSigned by: [redacted]	23-Feb-2021
SMT Statistician	Si	Date
[redacted]	DocuSigned by: [redacted]	25-Feb-2021
Program Statistician	Si	Date
[redacted]	DocuSigned by: [redacted]	23-Feb-2021
[redacted]	Si	Date

DocuSign Envelope ID: ECFD44FD-BFB1-4502-A562-95FFA3B107E2

VERSION HISTORY

SAP Version	Date	Primary Reasons for Amendment

DocuSign Envelope ID: ECFD44FD-BFB1-4502-A562-95FFA3B107E2

TABLE OF CONTENTS

DocuSign Envelope ID: ECFD44FD-BFB1-4502-A562-95FFA3B107E2

DocuSign Envelope ID: ECFD44FD-BFB1-4502-A562-95FFA3B107E2

LIST OF ABBREVIATIONS

Abbreviation	Definition
ACEND	Assessment of Caregiver Experience with Neuromuscular Disease
AE	adverse event
AESTDT	Start date of AE
CDC	Centers for disease control and prevention
CGIC	Clinical Global Impression of Change
CNS	central nervous system
CRF	case report form
CRO	contract research organization
CSF	cerebrospinal fluid
DHA	Directions for Handling and Administration
EAC	endpoint adjudication committee
ECG	electrocardiogram
GCP	Good Clinical Practice
HFMSE	Hammersmith Functional Motor Scale – Expanded
HINE	Hammersmith Infant Neurological Examination
HRQOL	health-related quality of life
ICF	informed consent form
ICH	International Council for Harmonisation
IDMC	independent data monitoring committee
IRT	interactive response technology
ITT	Intent-to-Treat
LP	lumbar puncture

DocuSign Envelope ID: ECFD44FD-BFB1-4502-A562-95FFA3B107E2

NOAEL	no-observed-adverse-effect level
NOEL	no-observed-effect level
PASA	Parent Assessment of Swallowing Ability
[redacted]	[redacted]
PedsQL	Pediatric Quality of Life Inventory™
[redacted]	[redacted]
[redacted]	[redacted]
QoL	quality-of-life
QTcF	corrected QT interval using Fridericia's formula
RSV	respiratory syncytial virus
RULM	Revised Upper Limb Module
SAE	serious adverse event
[redacted]	[redacted]
SMA	spinal muscular atrophy
SMN	survival motor neuron
SMN1	survival motor neuron 1
SMN2	survival motor neuron 2
SUSAR	suspected unexpected serious adverse reaction
[redacted]	[redacted]
TRTSDT	study treatment start date
US	United States
WHO	WHO World Health Organization

DocuSign Envelope ID: ECFD44FD-BFB1-4502-A562-95FFA3B107E2

1. Introduction

Nusinersen is an antisense oligonucleotide administered intrathecally via lumbar puncture (LP); it increases survival motor neuron (SMN) protein expression and significantly improves motor function in patients with spinal muscular atrophy (SMA). Nusinersen was approved for the treatment of SMA under the tradename Spinraza™ in the United States (US), European Union, and 15 other countries. The population for this study includes participants with infantile-onset and later-onset SMA.

2. Study Overview

2.1. Study Objectives and Endpoints

Study Primary Objective (Part A)

To examine the safety and tolerability of nusinersen administered intrathecally at higher doses to participants with SMA.

Study Primary Endpoint (Part A)

- Incidence of AEs, including SAEs
- Shifts from baseline in clinical laboratory parameters, ECGs, and vital signs
- Change in growth parameters
- Shifts from baseline in coagulation parameters (aPTT, PT, and INR)
- Change in urine total protein
- Change from baseline in neurological examination outcomes
- The proportion of participants with a post-baseline platelet count below the lower limit of normal on at least 2 consecutive measurements
- The proportion of participants with a post-baseline QTcF of> 500 msec and an increase from baseline to any post-baseline timepoint in QTcF of > 60 msec

Study Secondary Endpoint (Part A)

- Change from baseline in HFMSE score
- Change from baseline in RULM score
- Total number of new WHO motor milestones
- Change from baseline in ACEND
- Change from baseline in PedsQL
- Number and duration of hospitalizations
- CGIC (physician, caregiver) at Day 302
- Number of serious respiratory events
- Ventilator use
- Change in the PASA scale

DocuSign Envelope ID: ECFD44FD-BFB1-4502-A562-95FFA3B107E2

[redacted]

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

2.2. Study Design

This 3-part study will evaluate the efficacy and safety of a high dosing regimen of nusinersen in approximately 152 participants. The study will be conducted at approximately 65 sites globally. Following the completion of this study, all eligible participants may elect to enrol in a separate long-term extension study, pending study approval by ethics committees and the appropriate regulatory authorities.

This SAP is focusing only on Part A that is an open-label safety evaluation. Six participants with later-onset SMA who are 2 to ≤15 years of age, inclusive, at the signing of informed consent will receive 3 loading doses of 28mg of nusinersen on Days 1, 15, and 29, followed by 2 maintenance doses of 28 mg on Days 149 and 269.

2.3. Sample Size Considerations

A minimum of 6 participants with later-onset SMA will be enrolled in Part A to characterize the safety, tolerability, and [redacted] of a 28/28 mg dose of nusinersen (28-mg loading dose; 28-mg maintenance dose)

3. Definitions

3.1. Dates and Points of Reference

In order to distinguish nominal visit names from duration defined in days, visit names will be referred to as “Day 15”, “Day 29”, etc., and “15 days” or “29 days”, etc. will be used to define time intervals.

Summary statistics will be presented throughout. For continuous endpoints, summary statistics will generally include number of subjects with data, mean, standard deviation, median, minimum, and maximum. For categorical endpoints, summary statistics will generally include number of subjects dosed, number of subjects with data, and the percentage of those with data in each category. Frequency distributions will be presented as appropriate.

The statistical software, SAS® version 9.4 or above, will be used for all summaries and statistical analyses.

DocuSign Envelope ID: ECFD44FD-BFB1-4502-A562-95FFA3B107E2

Overall time on study will be defined as the total number of days a subject is known to be followed on study calculated as follows:

Overall time on Study = (Last date on study) – (Date of first dose) +1

The last date on study is defined as the date of the latest visit or evaluation or telephone contact, or time of death from all available data for a subject.

3.2. Study Treatment

Part A all participants will receive 3 loading doses of 28 mg (Days 1, 15, and 29) followed by 2 maintenance doses of 28 mg (Days 149 and 269).

3.3. Study Periods

N/A

3.4. Key Derived Variables

N/A.

3.5. Stratification Factors and Subgroup Variables

3.5.1. Stratification Factors

N/A.

3.5.2. Subgroup Variables

N/A

3.6. Analysis Sets

The Intent-to-Treat (ITT) Set is defined as all participants who receive at least a dose of nusinersen.

The Safety Analysis will be performed on the ITT Set.

[redacted]

DocuSign Envelope ID: ECFD44FD-BFB1-4502-A562-95FFA3B107E2

4. List of Planned Study Analyses

4.1. Interim Analysis

An Interim analysis may be performed to support publications or clinical decisions.

4.2. Primary Analysis

N/A

4.3. Final Analysis

The final analysis will be conducted after all Part A subjects have completed the study.

5. Statistical Methods for Planned Analyses

5.1. General Principles

Adverse Events

The first AE record occurs prior to the first dose, the second AE record occurs on or after the first dose, and the severity increases: Only the second record will be counted as treatment emergent.

The first AE record occurs prior to the first dose, the second AE record occurs on or after the first dose, and the severity decreases: Neither record will be counted as treatment emergent.

Both records occur on or after the first dose: If the AE severity on the second record is worse than the severity on the first record, then count both records as treatment emergent. But, if the severity improves, then only count the first record as treatment emergent.

Of note, when counting the total number of treatment-emergent events, events linked through change in severity will still be counted as separate events.

For events with missing start or stop dates, the following criteria will be used for the purpose of identifying treatment-emergent adverse events:

Denote the start date of Adverse Event as AESTDT and the study treatment start date as TRTSDT

- If both the start and stop dates for a particular event are missing, then the event is considered to have occurred on or after the first dose of study treatment;
- If the start date for a particular event is missing and the stop date/time falls after the first dose date/time, then the event is considered to have occurred on or after the first dose of study treatment;
- If the start time is missing and the start date is the same as the first dosing date, then the event is considered to have occurred on or after the first dose of study treatment;

DocuSign Envelope ID: ECFD44FD-BFB1-4502-A562-95FFA3B107E2

- If it cannot be determined whether or not an event has occurred on or after dosing due to a missing or partial date, then the event will be assumed to have occurred on or after the first dose for the purpose of identifying treatment-emergent adverse events.
- If AESTDT is completely missing or the year is missing, then impute AESTDT to TRTSTDT.
- If, in AESTDT, year is present and month/day are missing and year is equal to the year portion of TRTSTDT, then impute the month/day portion of AESTDT to the month/day portion of TRTSTDT.
- If, in AESTDT, year is present and month/day are missing and year is not equal to the year portion of TRTSTDT, then impute the month/day portion of AESTDT to January 01.
- Consider the situation in AESTDT where year and month are present with only day missing. If the year and month are the same as those for TRTSTDT, then impute day in AESTDT with day in TRTSTDT. Otherwise, impute the day in AESTDT with the first day of the month.

It is important to emphasize that the imputed date will not be used for calculations such as onset and duration of an adverse event. Due to the long half-life of nusinersen, analyses of treatment-emergent adverse events will include all events reported during the study.

Concomitant Medications and Procedure
In order to define concomitant therapies with missing start or stop dates, the following additional criteria will be used:

- If both the start and stop dates of a particular therapy were missing, that therapy is considered concomitant.
- If the start date of a therapy was missing and the stop date of that therapy fell on or after the date of dosing, that therapy is considered concomitant.
- If the start date of a therapy was prior to the date of dosing and the stop date of that therapy was missing and the therapy was listed as continuing, that therapy is considered concomitant.
- If the start date of a therapy was prior to the date of dosing and the stop date of that therapy was missing and the therapy was listed as not continuing, that therapy is considered concomitant, or
- If the start/stop date of a therapy is partial then where it is not possible to rule out that it was not taken concomitantly it will be considered concomitant

Denote the end date of medication as CMENDT. The medication is classified concomitant provided any of the following is NOT true:

DocuSign Envelope ID: ECFD44FD-BFB1-4502-A562-95FFA3B107E2

CMENDT is complete and CMENDT is less than TRTSDT

- Day of CMENDT missing and year/month of CMENDT is strictly before year-month of TRTSDT
- Month of CMENDT is missing and year of CMENDT is strictly before year of TRTSDT

5.2. Participant Accountability

The number of subjects who were screened, enrolled, dosed, and completed treatment and study, along with the reasons for discontinuing treatment and withdrawing from the study, will be presented.
Listings of those subjects who discontinued treatment and/or withdrew from the study and the reasons for discontinuation/withdrawal will be presented. Subjects who died during the study will be listed separately.

5.3. Demographic and Baseline Characteristics

Demography includes age categories, age at first dose, sex, ethnicity, and race. Medical history will be coded in MedDRA and the number and percentage of subjects with each history presented. SMA history includes age of symptom onset, age at SMA diagnosis, time from diagnosis to enrollment, number of copies of the SMN 1 and SMN2 genes, current motor function,

The 2000 CDC Growth Charts (ages 2 to <20 years) will be used to assess the weight change for older subjects. The National Center for Health Statistics provides a SAS macro which can be downloaded from their website [https://www.cdc.gov/growthcharts/clinical_ charts.htm] and this will be utilized to calculate the weight for age percentiles for each participant in the later onset population.

Subjects will be cross referenced with these files, given the age and sex of the subject to determine below which percentile they lie for each parameter.

Baseline quality of life and caregiver burden of SMA will be assessed by PedsQL and ACEND, respectively. Demographic, baseline disease characteristics, and baseline quality of life will be presented for the Safety Set.

5.4. Protocol Deviations

Protocol deviations identified during site monitoring will be captured in a protocol deviation log and categorized as major or minor deviations. These deviations will be listed. Major protocol deviations will be summarized.

DocuSign Envelope ID: ECFD44FD-BFB1-4502-A562-95FFA3B107E2

5.5. Study Treatment Exposure and Concomitant Medications

Exposure

The number of doses received will be summarized using frequency distributions. The amount of nusinersen received will be summarized using summary statistics.

Overall time on study definition see section 3.1.

Given the long half-life of nusinersen, participants are considered to be exposed from the time the first dose of nusinersen was administered (in or before the start of study) to the date of last visit or contact.

Concomitant Medications and Procedure

A concomitant therapy is any drug or substance administered between screening and the final study visit/telephone call. Participants should be instructed to contact their Investigators before taking any new medications, including nonprescription drugs and herbal preparations.

Throughout the study, Investigators or designated licensed physicians involved in the study may prescribe concomitant therapies (procedures) or treatments deemed necessary for AEs or to provide adequate supportive care.

Participants are prohibited from receiving other experimental or approved agents for the treatment of SMA, including gene therapy, biological agent, or device, during the study. This includes marketed agents at experimental doses that are being tested for the treatment of SMA (oral albuterol/salbutamol, riluzole, carnitine, sodium phenylbutyrate, valproate, hydroxyurea, etc.).

All concomitant medications will be coded using the World Health Organization drug dictionary (WHO Drug).

A concomitant procedure is any therapeutic intervention (e.g., surgery/biopsy, physical therapy) or diagnostic assessment (e.g., blood gas measurement, bacterial cultures) performed between Screening and the final study visit/telephone call. Concomitant procedures will be coded using MedDRA.

For the purposes of analysis, a concomitant therapy (including medication or procedure) is defined as any therapy that was taken or administered on or after the first injection of nusinersen. This includes therapies that were started prior to the initiation of injection of nusinersen if their use continued on or after the first injection of nusinersen. The number and percentage of subjects who were taking each type of concomitant medication at baseline and during the study will be presented. The number and percentage of subjects taking each type of ancillary procedure will be presented by preferred term.

DocuSign Envelope ID: ECFD44FD-BFB1-4502-A562-95FFA3B107E2

5.6. Efficacy Endpoints

No primary efficacy endpoints will be investigated.

The following secondary [redacted] efficacy assessments will be evaluated:

- HFMSE
- WHO Motor Milestones
- RULM
- ACEND
- PedsQL
- Hospitalizations
- CGIC
- Serious respiratory events
- Ventilator use
- PASA
- [redacted]

5.6.1. Primary Efficacy Endpoint

N/A

5.6.2. Secondary Efficacy Endpoints

Number of serious respiratory events

Respiratory events encompass a range of reported events and interventions. Adverse events coded into the SOC of respiratory, thoracic and mediastinal disorders will be defined as respiratory events and the subset of these which led to hospitalization will be identified using AE and hospitalization information together. Number of serious respiratory events will be listed for each subject.

Number and duration of hospitalizations

The number and duration of hospitalization will be listed for each subject.

Clinical Global Impression of Change

The CGIC rating scale was developed as a brief standalone assessment of the clinician's view of the patient's global functioning after initiating a study medication [Guy 1976]. The CGIC provides an overall clinician-determined summary measure that takes into account all available information, including knowledge of the patient's history, psychosocial circumstances, symptoms, behavior, and the impact of the symptoms on the patient's ability

DocuSign Envelope ID: ECFD44FD-BFB1-4502-A562-95FFA3B107E2

to function. The CGIC is administered by an experienced clinician who is familiar with the disease under study and the likely progression of treatment. The clinician makes a judgment about the total picture of the patient at each visit: the severity of the illness, the patient's level of distress and other aspects of impairment, and the impact of the illness on functioning. The CGIC is rated without regard to the clinician's belief that any clinical changes are or are not due to medication and without consideration of the etiology of the symptoms.

The CGIC is a 7-point scale that requires the clinician to assess how much the patient's illness has changed relative to a baseline state at the beginning of the intervention and is rated as follows: 1, very much improved; 2, much improved; 3, minimally improved; 4, no change; 5, minimally worse; 6, much worse; or 7, very much worse. CGIC should be assessed consistently by the same rater for each study participant. A separate CGIC assessment will be performed by the Investigator (Principal Investigator or Sub investigator) and caregiver. Adult participants who do not require a caregiver during the study visits will only have the CGIC assessment assessed by the Investigator.

The CGI score will be plotted over time for each subject and listed.

Hammersmith Functional Motor Scale – Expanded (HFMSE)

The HFMSE is a reliable and validated tool used to assess motor function in children with SMA. The scale was originally developed with 20 scored activities and was devised for use in children with Type II and Type III SMA with limited ambulation to give objective information on motor ability and clinical progression [Main 2003]. The expanded scale includes an additional module of 13 items developed to allow for the evaluation of ambulatory patients with SMA [O' Hagen 2007]

Each item is scored 0 (unable), 1 (performs with modification or adaptation) or 2 (able) and the total score is calculated by summing the 33 items and ranges from 0 to 66 with higher scores indicating greater motor function.

Baseline

The baseline for HFMSE is defined as the last non missing assessment prior to the first dose of study treatment.

Missing values

If 6 or fewer items are missing, then these items will be imputed to be 0 when summing all 33 items.

If greater than 6 items are missing, then the total score will be set to be missing; missing data will not be imputed.

DocuSign Envelope ID: ECFD44FD-BFB1-4502-A562-95FFA3B107E2

Total score and total score change from baseline will be plotted over time for each subject and listed.

Revised Upper Limb Module (RULM)

Participants with later-onset SMA will be evaluated using the Revised Upper Limb Module (RULM) [Mazzone 2016]. The RULM will continue to be performed should participants subsequently become ambulatory. The RULM is an outcome measure developed to assess upper limb functional abilities in patients with SMA, including young children, and patients with severe contractures in the lower limbs in whom the possibility of detecting functional changes, such as rolling or long sitting, is limited. This test consists of upper limb performance items that are reflective of activities of daily living (i.e., raise a can to the mouth as if drinking, take a coin and place it in a box, and remove the lid of a container). The RULM is quickly administered and has been evaluated in patients with SMA 2 to 52 years of age [Mazzone 2016]. The purpose of an upper limb scale for use in SMA is to assess the change that occurs in the motor performance of the upper limb over time. Motor performance in SMA is defined as a demonstrated ability to perform a skill under certain test conditions. This performance changes with disease progression and/or intervention (including surgery) and is based on the observed response on the day of the assessment. Motor performance will be impacted by muscle strength, contractures, and maturational development (puberty), and the RULM aims to incorporate the performance of the shoulder, elbow, wrist, and hand.

A derived total score will be calculated by summing the scores from these individual items. For each item, a score will be collected on the left and right side. A derived total score will be calculated by summing the scores from these 19 individual items and ranges from 0 if the subject fails all activities to 37 if the subject achieves all activities. If, for an individual item, a response is recorded for both the left and right side the highest score will be used in calculating the total.

Baseline

The baseline for RULM is defined as the last non missing assessment prior to the first dose of study treatment.

Missing values

If 3 or fewer items are missing, then these items will be imputed to be 0 when summing all items. If greater than 3 items are missing, then the total score will be set to be missing. Missing data will not be imputed.

Total score and total score change from baseline will be plotted over time for each subject and listed.

DocuSign Envelope ID: ECFD44FD-BFB1-4502-A562-95FFA3B107E2

WHO motor milestones

The WHO motor milestones are a set of six milestones in motor development, all of which would be expected to be attained by 24 months of age in healthy children. The individual milestones are: sitting without support, standing with assistance, hands and knees crawling, walking with assistance, standing alone and walking alone. The WHO reported [WHO Motor Development Study] that in about 90% of cases the order of attainment followed a fixed sequence for five of the milestones (namely, sitting without support, standing with assistance, walking with assistance, standing alone and walking alone) with only hands and knees crawling shifting between the earlier milestones.

The motor milestones will be assessed using the WHO motor milestone criteria [WHO Multicentre Growth Reference Study Group 2006]. As part of the assessment, the examiner records an overall rating of the subject's emotional state and then for each milestone one of the following four classifications:

- No (inability) – Child tried but failed to perform the milestone.
- No (refusal) – Child refused to perform despite being calm and alert.
- Yes – Child was able to perform the milestone.
- Unable to test – Could not be tested because of irritability, drowsiness or sickness.

Missing values

If for a milestone either 'No (refusal)' or 'Unable to test' are observed at a visit then the result will be first set to missing. Missing data will not be imputed.

Each milestone will be plotted in a binary way (0=not present/not achieved and 1=present /achieved) over time for each subject and a listing will be provided.

Quality-of-Life Questionnaires

QoL questionnaires include the Pediatric Quality of Life Inventory™(PedsQL) and Assessment of Caregiver Experience with Neuromuscular Disease (ACEND). QoL questionnaires will be collected on the days specified.

Pediatric Quality of Life Inventory (Generic Core Scales and Neuromuscular Module)

Participants in Part A will be evaluated using the PedsQL Measurement 4.0 Generic Core Scales and 3.0 Neuromuscular Module [Varni 1999]. This assessment has been validated by the American Spinal Muscular Atrophy Randomized Trials Group [Iannaccone 2009].

The PedsQL Measurement Model is a modular approach to measuring health-related quality of life (HRQOL) in children and adolescents. The PedsQL consists of brief, practical,

DocuSign Envelope ID: ECFD44FD-BFB1-4502-A562-95FFA3B107E2

generic core scales, as well as condition-specific modules for use in designated clinical populations. Pediatric self-report is measured in children and adolescents 5 to 18years of age, and parent proxy-report of child HRQOL is measured for children and adolescents 2 to 18 years of age. The PedsQL4.0 Generic Core Scales include an assessment of physical functioning, emotional functioning, social functioning, and school functioning. The PedsQL 3.0 Neuromuscular Module was designed to measure HRQOL dimensions specific to children 2 to 18years of age with neuromuscular disorders, including SMA.

Subjects in Part A can be between 2 and 15-year-old at screening.

The PedsQL parent questionnaire is collected for children in the following age categories: 2-4, 5-7, 8-12 and 13-18. Four dimensions are collected: Physical, Emotional, Social and School functioning and each item is scored on a 5-point ordinal scale (0= Never, 1 = Almost Never, 2= Sometimes, 3 = Often, 4 = Almost Always).

The PedsQL patient questionnaire is collected for children in the following categories: 5-7, 8-12 and 13-18. Similar dimensions and 5-point ordinal scale are used as for the parents but for subjects aged 5 to 18 years a 3-point ordinal scale is collected, omitting the response levels of 1 and 3.

In the neuromuscular module, one parent questionnaire is collected for all subjects irrespective of age with three dimensions: 'About my child's neuromuscular disease', Communication' and 'Family resources'. The same 5-point ordinal scale is collected for each question.

The patient neuromuscular disease questionnaire is collected for subjects in the following age categories: 5-7,8-12 and 13-18. The questionnaire for subjects aged 5-18 years uses the 3-point ordinal scale as above and has only one dimension - 'About my Neuromuscular disease'.

In scoring a dimension the first step is to reverse and linearly transform to a 0-100 scale (0 = 100, 1 =75, 2=50, 3= 25, 4= 0), so a higher score is indicative of a better health related quality of life.

A psychosocial health summary score, constructed from three dimensions, will be calculated as the sum of items over the number of items answered in the emotional, social and school functioning scales.

A total score will be calculated as the sum of all the items over the number of items answered on all the scales.

For the neuromuscular module, a score for each dimension and then total score will be calculated in the same manner, no health summary scores are evaluated.

Due to the age specific nature of these questionnaires, subjects aged 2-4 years would not be expected to complete the self-evaluation.

DocuSign Envelope ID: ECFD44FD-BFB1-4502-A562-95FFA3B107E2

Items scores, total scores and the change from baseline will be plotted over time for each subject and listed.

Baseline

The baseline for PedsQL is defined as the last non missing assessment prior to the first dose of study treatment.

Missing values

If greater than 50 percent of the items within a dimension are missing then the dimension score will not be computed, otherwise the mean score for the dimension will be calculated as the sum of items over the number of items answered.

Assessment of Caregiver Experience With Neuromuscular Disease (ACEND)

Parents/caregivers of participants will complete the ACEND questionnaire. The ACEND questionnaire will not be collected for adult participants who do not require a caregiver during the study visits. This assessment instrument has been designed to quantify the caregiver impact experienced by parents/caregivers of children affected with severe neuromuscular diseases, including children with SMA [Matsumoto 2011]. ACEND includes domains assessing physical impact (including feeding/grooming/dressing, sitting/play, transfers, and mobility) and general caregiver impact (including time, emotion, and finance). and each domain comprises several items.

The total score for a domain with n items, each item assessed on ordinal scale from 1 to z, is derived as follows: 100 multiplied by (Mean of the n items in the domain -1) divided by (z-1).

This total score will be on a scale of 0 to 100 with a higher score indicating a greater impact on the caregiver.

All items score, total score and the change from baseline will be plotted over time for each subject and listed.

Baseline

The baseline for ACEND is defined as the last non missing assessment prior to the first dose of study treatment.

Missing values

At least two items for the time domain and one item for the remaining domains need to be non-missing for a total to be calculated; else the total score will be set to be missing.

DocuSign Envelope ID: ECFD44FD-BFB1-4502-A562-95FFA3B107E2

Parent Assessment of Swallowing Ability

Will be assessed in Part A at the timepoints specified in the Schedule of Activities using the PASA questionnaire. Caregivers will be asked a series of questions regarding the mealtime behavior of the participant. The PASA questionnaire was developed by a Biogen team in order to assess the signs and symptoms of dysphagia. This questionnaire consists of 33 items across 4 domains that cover general feeding, drinking liquids, eating solid foods, and assessment of swallowing concerns. The first 3 of these domains are generally assessed with 5 levels of response (Never, Rarely, Sometimes, Often, and Always), although 2 items are assessed with a simple "Yes"/"No" answer. In the final domain, the assessment of swallowing concerns has 4 levels of response: Strongly Agree, Agree, Disagree, and Strongly Disagree. In answering each item, the caregiver is directed to consider the previous 7 days.

All of the item results will be plotted over time for each subject and listed.

Ventilator Use

The participant's ventilator use will be collected at every study visit. If ventilation is used daily, the average number of hours per day for the past 7 days will be recorded.

The ventilation use and number of hours collected and will be listed.

[redacted]

DocuSign Envelope ID: ECFD44FD-BFB1-4502-A562-95FFA3B107E2

[redacted]

5.7. Primary Safety Endpoints

Analyses of safety data will include adverse events and serious adverse events, laboratory data, ECGs, growth parameters, vital signs, and neurological examinations. Worsening or new findings noted from the physical or neurological examinations will be reported as AEs as appropriate.

Adverse events

All adverse events (AEs) will be analyzed based on the principle of treatment emergence. An adverse event will be regarded as treatment-emergent if it was present prior to receiving the

DocuSign Envelope ID: ECFD44FD-BFB1-4502-A562-95FFA3B107E2

first dose of nusinersen and subsequently worsened in severity, or was not present prior to receiving the first dose of nusinersen and subsequently appeared.

Adverse events will be coded using the MedDRA dictionary. This coding system provides more than five levels to classify adverse events. In general, adverse events will be presented by system organ class and preferred terms but other classifications may be used if warranted.

The incidence and frequency (event count) of treatment-emergent adverse events will be summarized. A subject having the same adverse event more than once will be counted only once in the incidence for that adverse event; multiple occurrences of the same adverse event for the same subject will all be counted in the frequency for that adverse event. The summary tables will include incidence estimates for overall system organ class as well as for preferred terms within each system organ class. Incidence will be presented by decreasing order by system organ class and by decreasing order by preferred term within each system organ class. The most common adverse events, i.e., those that occurred in at least 3 subjects, will be presented. Upon examination of the actual data, different cut-offs may be used if it is deemed more appropriate.

Clinical laboratory data

The following clinical laboratory parameters are to be assessed:

- Hematology: red blood cells, hemoglobin, hematocrit, platelets, white blood cells, white blood cell differential, neutrophils, eosinophils, basophils, lymphocytes and monocytes.
- Blood chemistry: sodium, potassium chloride, total protein, albumin, calcium, phosphorus bicarbonate, glucose, blood urea nitrogen, creatinine, cystatin C, total serum bilirubin (direct and indirect), alkaline phosphatase, aspartate aminotransferase; alanine aminotransferase, creatine phosphokinase, creatine kinase and gamma glutamyl transferase.
- Urinalysis: specific gravity, pH, protein, glucose, ketones, bilirubin, red blood cells, white blood cells, epithelial cells, bacteria, casts, crystals, Cystatin C (for subjects with Screening weight of ≥3 kg) and urine total protein assessed by local laboratories
- Coagulation: aPTT, PT, INR

For a parameter if the local and central result are available with the same date and time then only the central analysis result will be considered for presentations by visit. In a situation where two or more results have the same date and time and are both central (or both local) then we will check with the Safety physician if it would be more appropriate or take the highest or lowest value.

As described below, laboratory data will be examined using an analysis of "shifts" and all available data will be considered for these irrespective if collected centrally or locally. Each subject's laboratory values will be classified according to whether the test result is "low" (i.e., below the lower limit of normal [LLN]), "normal" (within the normal range), or "high" (i.e., above the upper limit of normal [ULN]). If a subject is missing a baseline value but had

DocuSign Envelope ID: ECFD44FD-BFB1-4502-A562-95FFA3B107E2

a post-baseline value, then the baseline assessment is labeled as "unknown". Likewise, if a subject had a baseline value but had no post-baseline values, then the minimum and maximum are labeled as "unknown".

All the parameter above will be listed (including change and shift from baseline) and plotted for each subject. Also, the participants with a post-baseline platelet count below the lower limit of normal on at least 2 consecutive measurements will be listed.

Baseline

The baseline for all the laboratory parameters is defined as the last non missing assessment prior to the first dose of study treatment.

CSF local laboratory samples.

Cell count, protein, and glucose from the CSF local laboratory sample will be listed and plotted for each subject.

ECG

ECGs are to be recorded at Days 1, 15, 29, 149, 269 and 302/ET, pre-dosing and post-dosing at various time points on dosing days and end pre-dosing at 302/ET. The ECGs are assessed at a central reading laboratory and the results provided as external vendor data. On the eCRF the investigator interpretation is also collected

ECG listings will be presented for all subjects and for subjects with post-baseline QTcF of > 500 msec and increase from baseline >60 msec.

Baseline

The baseline for ECG is defined as the assessment prior to the first dose of study treatment.

Growth parameters

Growth parameters comprise height/length, ulnar length, weight for age percentile and length for age percentile will be listed.

LP opening pressure.

LP opening pressure results will be plotted over time for each subject and a listed.

Vital signs

Vital signs are to be measured at Screening, pre-dosing and at various time points post-dosing on dosing days. At each of these times, temperature, heart rate, respiratory rate, systolic and diastolic blood pressure, and pulse oximetry awake will be measured.

Vital sign will be listed.

DocuSign Envelope ID: ECFD44FD-BFB1-4502-A562-95FFA3B107E2

Baseline

The baseline for vital sign is defined as the last non missing assessment prior to the first dose of study treatment.

Neurological exams

Standard neurological examinations, which include assessments of mental status, level of consciousness, sensory function, motor function, cranial nerve function, and reflexes, will be conducted.

The neurological items comprise of mental status, level of consciousness, sensory function, motor function, cranial nerve function, and reflexes will be assessed. These assessments will be done at screening, pre-dosing and post-dosing.

The result collected for the majority of the tests is 'normal' or 'abnormal', except for the assessment of sensory function which is reported as 'present' or 'absent', and the assessment of reflexes which is captured on an ordinal scale. For each abnormal test result, it is recorded whether or not it is secondary to SMA.
All the neurological items result will be plotted over time for each subject and listed.

[redacted]

DocuSign Envelope ID: ECFD44FD-BFB1-4502-A562-95FFA3B107E2

[redacted]

DocuSign Envelope ID: ECFD44FD-BFB1-4502-A562-95FFA3B107E2

[redacted]

6. Changes from Protocol-Specified Analyses

N/A

7. Summary of Changes from the Previous Version of the SAP

N/A

8. References

CDC growth scale https://www.cdc.gov/growthcharts/clinical_charts.htm

Iannaccone ST, Hynan LS, Morton A, et al. The PedsQL in pediatric patients with Spinal Muscular Atrophy: feasibility, reliability, and validity of the Pediatric Quality of Life Inventory Generic Core Scales and Neuromuscular Module. Neuromuscul Disord. 2009;19(12):805-12.

Guy W. ECDEU Assessment Manual for Psychopharmacology. Rockville, MD: U.S. Department of Health, Education, and Welfare; 1976.

Main M, Kairon H, Mercuri E, et al. The Hammersmith functional motor scale for children with spinal muscular atrophy: a scale to test ability and monitor progress in children with limited ambulation. Eur J Paediatr Neurol. 2003;7(4):155-9.

Matsumoto H, Clayton-Krasinski DA, Klinge SA, et al. Development and initial validation of the assessment of caregiver experience with neuromuscular disease. J Pediatr Orthop. 2011;31(3):284-92.

Mazzone ES, Mayhew A, Montes J, et al. Revised upper limb module for spinal muscular atrophy: Development of a new module. Muscle Nerve. 2016 Epub 2016/10/04.

DocuSign Envelope ID: ECFD44FD-BFB1-4502-A562-95FFA3B107E2

[redacted]

[redacted]

O'Hagen JM, Glanzman AM, McDermott MP, et al. An expanded version of the Hammersmith Functional Motor Scale for SMA II and III patients. Neuromuscul Disord. 2007;17(9-10):693-7.

[redacted]

Varni JW, Seid M, Rode CA. The PedsQL: measurement model for the pediatric quality of life inventory. Med Care. 1999;37(2):126-39.

WHO Multicentre Growth Reference Study Group. Assessment of sex differences and heterogeneity in motor milestone attainment among populations in the WHO Multicentre Growth Reference Study. Acta Paediatr Suppl. 2006a; 450:66-75.

WHO Motor Development Study: Windows of achievement for six gross motor development milestones, Acta Pædiatrica, 2006; Suppl 450: 86/95.

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

DocuSign Envelope ID: ECFD44FD-BFB1-4502-A562-95FFA3B107E2

APPENDICES

N/A

Certificate Of Completion

Envelope Id: ECFD44FDBFB14502A56295FFA3B107E2
Status: Completed
Subject: Please DocuSign: 232SM203 Part A Statistical Analysis V1.0.pdf
Vendor Name:
Source Envelope:
Document Pages: 29
Signatures: 3
Certificate Pages: 5
Initials: 0
AutoNav: Enabled
EnvelopeId Stamping: Enabled
Time Zone: (UTC-05:00) Eastern Time (US & Canada)

Envelope Originator:
[redacted]
225 Binney St
Cambridge, MA 02142
[redacted]
IP Address: [redacted]

Record Tracking

Status: Original
23 February 2021 | 09:31
Holder: [redacted]
Location: DocuSign

Signer Events	Signature	Timestamp
[redacted] Principal Biostatistician Biogen Security Level: Email, Account Authentication (None)	DocuSigned by: [redacted] Signature Adoption: Pre-selected Style Using IP Address: [redacted]	Sent: 23 February 2021 | 09:40 Viewed: 23 February 2021 | 09:44 Signed: 23 February 2021 | 09:44
Electronic Record and Signature Disclosure: Not Offered via DocuSign		
[redacted] Biogen Security Level: Email, Account Authentication (None)	DocuSigned by: [redacted] Signature Adoption: Pre-selected Style Using IP Address: [redacted]	Sent: 23 February 2021 | 09:40 Viewed: 25 February 2021 | 08:28 Signed: 25 February 2021 | 08:29
Electronic Record and Signature Disclosure: Not Offered via DocuSign		
[redacted] Medical Director Security Level: Email, Account Authentication (None)	DocuSigned by: [redacted] Signature Adoption: Pre-selected Style Using IP Address: [redacted]	Sent: 23 February 2021 | 09:40 Viewed: 23 February 2021 | 11:08 Signed: 23 February 2021 | 11:09
Electronic Record and Signature Disclosure: Accepted: 23 February 2021 | 11:08 ID: 2bad1685-17cd-4ba8-88b7-1a9d74a30a77		

In Person Signer Events	Signature	Timestamp

Editor Delivery Events	Status	Timestamp

Agent Delivery Events	Status	Timestamp

Intermediary Delivery Events	Status	Timestamp

Certified Delivery Events	Status	Timestamp

Carbon Copy Events	Status	Timestamp

Witness Events	Signature	Timestamp

Notary Events	Signature	Timestamp

Envelope Summary Events	Status	Timestamps
Envelope Sent	Hashed/Encrypted	23 February 2021 | 09:40
Certified Delivered	Security Checked	23 February 2021 | 11:08
Signing Complete	Security Checked	23 February 2021 | 11:09
Completed	Security Checked	25 February 2021 | 08:29

Payment Events	Status	Timestamps

Electronic Record and Signature Disclosure

ELECTRONIC RECORD AND SIGNATURE DISCLOSURE

From time to time, Biogen (we, us or Company) may be required by law to provide to you certain written notices or disclosures. Described below are the terms and conditions for providing to you such notices and disclosures electronically through your DocuSign, Inc. (DocuSign) Express user account. Please read the information below carefully and thoroughly, and if you can access this information electronically to your satisfaction and agree to these terms and conditions, please confirm your agreement by clicking the 'I agree' button at the bottom of this document.

Getting paper copies

At any time, you may request from us a paper copy of any record provided or made available electronically to you by us. For such copies, as long as you are an authorized user of the DocuSign system you will have the ability to download and print any documents we send to you through your DocuSign user account for a limited period of time (usually 30 days) after such documents are first sent to you. After such time, if you wish for us to send you paper copies of any such documents from our office to you, you will be charged a $0.00 per-page fee. You may request delivery of such paper copies from us by following the procedure described below.

Withdrawing your consent

If you decide to receive notices and disclosures from us electronically, you may at any time change your mind and tell us that thereafter you want to receive required notices and disclosures only in paper format. How you must inform us of your decision to receive future notices and disclosure in paper format and withdraw your consent to receive notices and disclosures electronically is described below.

Consequences of changing your mind

If you elect to receive required notices and disclosures only in paper format, it will slow the speed at which we can complete certain steps in transactions with you and delivering services to you because we will need first to send the required notices or disclosures to you in paper format, and then wait until we receive back from you your acknowledgment of your receipt of such paper notices or disclosures. To indicate to us that you are changing your mind, you must withdraw your consent using the DocuSign 'Withdraw Consent' form on the signing page of your DocuSign account. This will indicate to us that you have withdrawn your consent to receive required notices and disclosures electronically from us and you will no longer be able to use your DocuSign Express user account to receive required notices and consents electronically from us or to sign electronically documents from us.

All notices and disclosures will be sent to you electronically

Unless you tell us otherwise in accordance with the procedures described herein, we will provide electronically to you through your DocuSign user account all required notices, disclosures, authorizations, acknowledgements, and other documents that are required to be provided or made available to you during the course of our relationship with you. To reduce the chance of you inadvertently not receiving any notice or disclosure, we prefer to provide all of the required notices and disclosures to you by the same method and to the same address that you have given us. Thus, you can receive all the disclosures and notices electronically or in paper format through the paper mail delivery system. If you do not agree with this process, please let us know as described below. Please also see the paragraph immediately above that describes the consequences of your electing not to receive delivery of the notices and disclosures electronically from us.

How to contact Biogen:
You may contact us to let us know of your changes as to how we may contact you electronically, to request paper copies of certain information from us, and to withdraw your prior consent to receive notices and disclosures electronically as follows:
To contact us by email send messages to: [redacted]

To advise Biogen of your new e-mail address
To let us know of a change in your e-mail address where we should send notices and disclosures electronically to you, you must send an email message to us at [redacted] and in the body of such request you must state: your previous e-mail address, your new e-mail address. We do not require any other information from you to change your email address..
In addition, you must notify DocuSign, Inc to arrange for your new email address to be reflected in your DocuSign account by following the process for changing e-mail in DocuSign.

To request paper copies from Biogen
To request delivery from us of paper copies of the notices and disclosures previously provided by us to you electronically, you must send us an e-mail to [redacted] and in the body of such request you must state your e-mail address, full name, Postal address, and telephone number. We will bill you for any fees at that time, if any.

To withdraw your consent with Biogen
To inform us that you no longer want to receive future notices and disclosures in electronic format you may:

i. decline to sign a document from within your DocuSign account, and on the subsequent page, select the check-box indicating you wish to withdraw your consent, or you may;
ii. send us an e-mail to [redacted] and in the body of such request you must state your e-mail, full name, IS Postal Address, telephone number, and account number. We do not need any other information from you to withdraw consent.. The consequences of your withdrawing consent for online documents will be that transactions may take a longer time to process..

Required hardware and software

Operating Systems:	Windows2000? or WindowsXP?
Browsers (for SENDERS):	Internet Explorer 6.0? or above
Browsers (for SIGNERS):	Internet Explorer 6.0?, Mozilla FireFox 1.0, NetScape 7.2 (or above)
Email:	Access to a valid email account
Screen Resolution:	800 x 600 minimum
Enabled Security Settings:	•Allow per session cookies •Users accessing the internet behind a Proxy Server must enable HTTP 1.1 settings via proxy connection

** These minimum requirements are subject to change. If these requirements change, we will provide you with an email message at the email address we have on file for you at that time providing you with the revised hardware and software requirements, at which time you will have the right to withdraw your consent.

Acknowledging your access and consent to receive materials electronically
To confirm to us that you can access this information electronically, which will be similar to other electronic notices and disclosures that we will provide to you, please verify that you were able to read this electronic disclosure and that you also were able to print on paper or electronically save this page for your future reference and access or that you were able to e-mail this disclosure and consent to an address where you will be able to print on paper or save it for your future reference and access. Further, if you consent to receiving notices and disclosures exclusively in electronic format on the terms and conditions described above, please let us know by clicking the 'I agree' button below.
By checking the 'I Agree' box, I confirm that:

- I can access and read this Electronic CONSENT TO ELECTRONIC RECEIPT OF ELECTRONIC RECORD AND SIGNATURE DISCLOSURES document; and

- I can print on paper the disclosure or save or send the disclosure to a place where I can print it, for future reference and access; and

- Until or unless I notify Biogen as described above, I consent to receive from exclusively through electronic means all notices, disclosures, authorizations, acknowledgements, and other documents that are required to be provided or made available to me by Biogen during the course of my relationship with you.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

PART B INFANTILE-ONSET SMA

STATISTICAL ANALYSIS PLAN

Version No.: 3.0

Date: 19Jul2024

Author: [redacted]

Study Title:

Escalating Dose and Randomized, Controlled Study of Nusinersen (BIIB058) in Participants With Spinal Muscular Atrophy

Name of Study Treatment: Nusinersen

Protocol No.: 232SM203

Study Phase: 2/3

Confidential Information

The information contained in this Statistical Analysis Plan (SAP) is confidential and may not be reproduced or otherwise disseminated without the approval of Biogen, Inc.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

APPROVAL

This document has been reviewed and approved by:		
[redacted]	DocuSigned by: [redacted]	19-Jul-2024
SMT Statistician	Signature	Date
[redacted]	DocuSigned by: [redacted]	19-Jul-2024
Program Statistician	Signature	Date
[redacted]	DocuSigned by: [redacted]	19-Jul-2024
SMT Medical Director	Signature	Date

VERSION HISTORY

SAP Version	Date	Primary Reasons for Amendment
Version 2	18July2024	Updated based on FDA comments

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

Version 3	19July2024 (prior to release of treatment codes)	Updated based on FDA comment

TABLE OF CONTENTS

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

LIST OF ABBREVIATIONS

Abbreviation	Definition
[redacted]	[redacted]
[redacted]	[redacted]
aPTT	Activated Partial Thromboplastin Time
ACEND Assessment of Caregiver Experience with Neuromuscular Disease	Assessment of Caregiver Experience with Neuromuscular Disease
[redacted]	[redacted]
AE	Adverse event
ALT	Alanine Aminotransferase (SGPT)
AST	Aspartate Aminotransferase (SGOT)

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

[redacted]	[redacted]
BLQ	Below limit of quantification
BUN	Blood Urea Nitrogen
CGIC	Clinical Global Impression of Change
CGIC	Clinical Global Impression of Change
CHOP INTEND	Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders
[redacted]	[redacted]
CO2	Carbon dioxide
COVID-19	Coronavirus disease 2019
CRF	Case report form
CSF	Cerebrospinal fluid
CV	Coefficient of variance
[redacted]	[redacted]
EAC	Endpoint adjudication committee
ECG	Electrocardiogram
GGT	Gamma-Glutamyl Transferase
HFMSE	Hammersmith Functional Motor Scale – Expanded
HINE	Hammersmith Infant Neurological Examination
HRQOL	Health-related quality of life
ICF	Informed consent form
ICH	International Council for Harmonization
INR	International Normalized Ratio
IDMC	Independent data monitoring committee
IRT	Interactive response technology
ITT	Intent-to-Treat
JRT	Joint Rank Test
LLOQ	Lower Limit of Quantification

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

LP	Lumbar puncture
MI	Multiple imputation
NF-L	Neurofilament Light chain
PASA	Parent Assessment of Swallowing Ability
PD	Pharmacodynamic(s)
PT	Prothrombin Time
PedsQL	Pediatric Quality of Life Inventory™
PK	Pharmacokinetic(s)
PPS	Per protocol Set
QoL	Quality-of-life
QTcF	Corrected QT interval using Fridericia's formula
RULM	Revised Upper Limb Module
SAE	Serious adverse event
SD	Standard deviation
SE	Standard error
SMA	Spinal muscular atrophy
SMN	Survival motor neuron
SMN1	Survival motor neuron 1
SMN2	Survival motor neuron 2
SOC	System Organ Class
[redacted]	[redacted]
ULN	Upper Limit of Normal
US	United States
WHO	World Health Organization

1. Introduction

This Statistical Analysis Plan is based on Version 7 of the protocol, dated 14Jun2024. All references to the protocol refer to Version 7.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

Nusinersen is an antisense oligonucleotide administered intrathecally via lumbar puncture (LP); it increases survival motor neuron (SMN) protein expression and significantly improves motor function in patients with spinal muscular atrophy (SMA). Nusinersen was approved for the treatment of SMA under the tradename Spinraza™ in the United States (US), European Union, and other countries. The population for this study includes participants with infantile-onset and later-onset SMA.

Efficacy and safety results across the nusinersen clinical development program have demonstrated an overall positive benefit-risk profile of nusinersen across a broad range of SMA phenotypes and patient populations. Nusinersen is approved in the US, Europe, and other countries and regions for the treatment of SMA in paediatric and adult patients at a recommended dosage of 12 mg administered in 3 loading doses at 14-day intervals, a fourth loading dose 30 days after the third dose, and maintenance doses every 4 months thereafter in most countries and regions. Pharmacokinetic (PK) and pharmacodynamic (PD) analyses indicate that nusinersen drug exposure higher than that achieved with 12 mg in patients with SMA may produce an even greater benefit in motor function. Additionally, PK modelling and simulations identified dosing regimens that achieve higher drug exposure more rapidly. Therefore, this study is being conducted to investigate the efficacy, safety, tolerability, and PK of a 50/28-mg dose of nusinersen (50 mg loading dose/28 mg maintenance dose) and a dosing regimen targeted to achieve higher drug exposure more rapidly. This study will be conducted in participants with genetically confirmed SMA.

This is a 3-part study Part A, Part B (infantile-onset and later-onset), and Part C.

Part A was an open label safety evaluation in which later-onset SMA subjects received 3 loading doses of 28 mg of nusinersen and 2 maintenance doses of 28 mg. All 6 participants enrolled in Part A completed the study.

Part B was double-blind, active-controlled controlled study designed to evaluate the proposed higher dosing regimen and included infantile and later-onset SMA participants.

Part C was an open label safety evaluation in which infantile and later-onset SMA of the participants, who were on the currently approved dose of nusinersen (12-mg maintenance for at least 1 year after the initiation of treatment) received higher dosing regimen via the administration of a single bolus dose of 50 mg of nusinersen (which should be administered 4 months ± 14 days after their most recent maintenance dose of 12 mg), with maintenance dosing at 28 mg thereafter. All 40 participants enrolled in Part C completed.

maintenance dose of 12 mg), with maintenance dosing at 28 mg thereafter.

This SAP will cover the planned analysis of Part B infantile-onset population, a separate SAP will cover the later-onset population.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

2. Study Overview

2.1. Study Objectives and Endpoints

This analysis plan is for Part B infantile-onset. Separate SAPs will cover the planned analyses for other parts of 232SM203.

Primary Objective

To examine the clinical efficacy of nusinersen administered intrathecally at higher doses to participants with SMA in the 232SM203 50/28 mg Group compared to the CS3B Matched Sham Control Group, as measured by change in CHOP INTEND total score.

Primary Endpoint

- Change from baseline to Day 183 in CHOP INTEND total score, accounting for mortality/dropout using the joint-rank test (comparison of higher dose to matched sham control)

Secondary Objectives

To examine the clinical efficacy of nusinersen administered intrathecally at higher doses to participants with SMA in the 232SM203 50/28 mg group compared to:

- CS3B Matched Sham Control Group
- AND 232SM203 nusinersen 12 mg Group

Secondary Endpoints

In a change from protocol V7.0, at the time of finalizing this Statistical Analysis Plan it was decided to change the timepoint for the testing of plasma NF-L for the 50/28mg versus 12mg comparison from Day 29 to Day 64.

Endpoint	**Comparison:** 50/28-mg Group versus
Proportion of Hammersmith Infant Neurological Examination (HINE) Section 2 motor milestone responders at Day 183	Matched Sham Set
Change from baseline to Day 183 in HINE Section 2 motor milestones total score accounting for mortality/dropout using the joint-rank test	Matched Sham Set

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

Change from baseline to Day 183 in plasma concentration of NF-L.	Matched Sham Set
Change from baseline to Day 302 in CHOP INTEND total score	12 mg nusinersen
Change from baseline to Day 302 in HINE Section 2 motor milestones total score	12 mg nusinersen
Change from baseline to Day 64 in plasma concentration of NF-L	12 mg nusinersen
Time to death or permanent ventilation (tracheostomy or ≥ 16 hours of ventilation/day continuously for > 21 days in the absence of an acute reversible event)	Matched Sham Set
Time to death (overall survival)	Matched Sham Set
Time to death or permanent ventilation (tracheostomy or ≥ 16 hours of ventilation/day continuously for > 21 days in the absence of an acute reversible event)	12 mg nusinersen
Time to death (overall survival)	12 mg nusinersen

Secondary Safety Objectives

To examine the safety and tolerability of nusinersen administered intrathecally at higher doses to participants with SMA.

Secondary Safety Endpoints

- Incidence of AEs, including SAEs
- Shifts from baseline in clinical laboratory parameters, ECGs, and vital signs
- Change in growth parameters
- Shifts from baseline in coagulation parameters (aPTT, PT, and INR)
- Change in urine total protein
- Change from baseline in neurological examination outcomes
- The proportion of participants with a post baseline platelet count below the lower limit of normal on at least 2 consecutive measurements
- The proportion of participants with a post baseline QTcF of > 500 msec and an increase from baseline to any post baseline timepoint in QTcF of > 60 msec

Additional Secondary Objectives

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

To examine the effect of nusinersen administered intrathecally at higher doses compared to the currently approved dose in participants with SMA

Secondary Endpoints

- Number and duration of hospitalizations
- Clinical Global Impression of Change (CGIC) [physician, caregiver] at Day 302
- Number of serious respiratory events
- Proportion of time on ventilation
- Ventilator use
- Change in the Parent Assessment of Swallowing Ability (PASA) scale
- Change from baseline in CSF concentration of NF-L

[redacted]

- [redacted]

2.2. Study Design

This 3-part study will evaluate the efficacy and safety of a high dosing regimen of nusinersen in approximately 145 participants. The study will be conducted at approximately 65 sites globally. Following the completion of this study, all eligible participants may elect to enrol in a separate long-term extension study (232SM302), pending study approval by ethics committees and the appropriate regulatory authorities. In regard to participants rolling over from Part B, this will be done without unblinding the participant's treatment group.

Part B will consist of a pivotal, double-blind, active-controlled treatment period with either 2 or 4 loading doses of nusinersen (for the 50/28 mg Group and Control Group, respectively) administered intrathecally followed by maintenance doses approximately every 4 months thereafter. Approximately 75 participants with infantile-onset SMA will be randomized in a 1:2 ratio to receive either the currently approved dosing regimen or a 50/28 mg dosing regimen.

In the Control Group, up to a total of 25 participants will receive 4 loading doses of 12 mg of nusinersen administered intrathecally (Days 1, 15, 29, and 64), followed by maintenance doses of 12 mg on Days 183 and 279 (and a sham procedure on Day 135).

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

In the 50/28 mg Group, up to a total of 50 participants will receive 2 loading doses of 50 mg of nusinersen administered intrathecally (Days 1 and 15) and 2 sham procedures on Days 29 and 64, followed by maintenance doses of 28 mg on Days 135 and 279 (and a sham procedure on Day 183). In order to maintain blinding sham procedures will be used as shown in the table below.

Arm	Dose 1	Dose 2	Dose 3	Dose 4	Dose 5	Dose 6	Dose 7
50/28 mg Group	D1 (50 mg)	D15 (50 mg)	D29 (sham)	D64 (sham)	D135 (28 mg)	D183 (sham)	D279 (28 mg)
Control (12/12 mg) Group	D1 (12 mg)	D15 (12 mg)	D29 (12 mg)	D64 (12 mg)	D135 (sham)	D183 (12 mg)	D279 (12 mg)

Participants will remain at the clinic for at least 24 hours after each study treatment administration or sham procedure.

Randomization in Part B will be performed using interactive response technology (IRT) and the randomization will be stratified by disease duration as follows:

≤ 12 weeks and

> 12 weeks (time from age at symptom onset to age at informed consent)

Blinded safety data will be reviewed on an ongoing basis by the Sponsor and the Medical Monitor. Unblinded safety data will be reviewed on an ongoing basis by an Independent Data Monitoring Committee (IDMC).

Time to death or permanent ventilation will be determined in a blinded fashion by a central, independent Endpoint Adjudication Committee (EAC). Procedures for reviewing and adjudicating events are described in the charter that governs the operation of the EAC.

2.3. Sample Size Considerations

The justification for the sample size for the infantile-onset SMA population in Part B is detailed as follows.

For the purposes of sample size determination, it is assumed that the higher dose will achieve an additional 4.5 improvement in CHOP INTEND and additional 3% in survival over the 12 mg dose. For the infantile-onset SMA population in Part B, a sample size of approximately 50 participants in the 50/28 mg Group and 20 sham subjects from study CS3B will provide at least approximately 99% power for the primary endpoint to detect an improvement of 24 points

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

on CHOP INTEND and 23% survival rate benefit (compared to that observed in Study CS3B participants receiving sham control) at Day 183 based on the joint rank test at a 2 sided significance level of 0.05. This power calculation is based on simulations using data generated from a joint model of survival and functional change. The model used a difference of 24 points for the Day 183 change from baseline in CHOP INTEND total score (50/28 mg Group – Study CS3B Sham Control Group) and a population standard deviation of 8.8 for change from baseline.

2.4. Historical Data

CS3B Study

CS3B was a phase 3, randomized, double-blind, sham-procedure controlled study to assess the clinical efficacy and safety of 12 mg nusinersen administered intrathecally in participants with infantile-onset SMA. 122 participants were randomized 2:1 to receive a scaled equivalent 12 mg dose of nusinersen or underwent a sham procedure as control, respectively.

Key eligibility criteria were SMN2 copy number = 2, onset of SMA symptoms <= 180 days of age and age at screening <=210 days. The study had two primary endpoints:

1) Proportion of HINE Section 2 motor milestone responders and
2) Time to death or permanent ventilation.

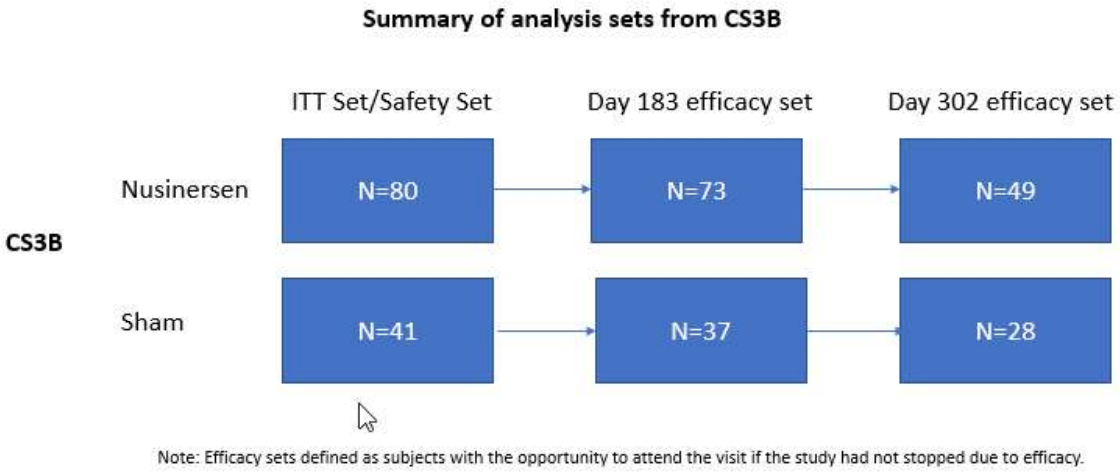
A key secondary endpoint was change in CHOP INTEND.

The first participant in CS3B was dosed in August 2014 and the last participant last visit was November 2016. The study was stopped early for efficacy based on the results of an interim analysis and the final study results were published (Finkel et al. 2017).

Figure Showing Key Analysis Sets From CS3B

The figure below shows the number of participants in the ITT/Safety set and efficacy sets for the CS3B.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365



Summary of How Historical Data Will be Used in the Analysis

The enrollment criteria of the Part B infantile-onset in 232SM203 has been designed to be similar to CS3B study (external data). In general, consistent inclusion/exclusion criteria have been utilized and efficacy assessments at the same timepoints have been included.

As part of the analysis of the 232SM203 study the external data will be utilized for the following purposes:

- For the primary and secondary endpoints comparing to sham, the higher dose participants from 232SM203 will be compared to matched sham control participants from CS3B. Comparability will be assessed for 232SM203 compared to CS3B sham based on key covariates and a matching algorithm will be utilized to select at least 20 CS3B sham control subjects. For the remaining secondary endpoints, the high dose participants will be compared to 12 mg dose using the 232SM203 ITT population.
- [redacted]
- Use of the group of matched control subjects described in the first bullet for summaries of safety endpoints. This is in order to provide a reference group of participants with similar disease who have not undergone lumbar puncture SMA.

3. Definitions

3.1. Dates and Points of Reference

- Study Day 1: the date of the first dose of study treatment in 232SM203

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

- Study Day
  - For a date on or after Study Day 1

    Study Day = (Date of Interest) – (Study Day 1) + 1
  - For a date before Study Day 1

    Study Day = (Date of Interest) – (Study Day 1)

In order to distinguish nominal visit names from duration defined in days, visit names will be referred to as "Day 15", "Day 29", etc., and "15 days" or "29 days", etc. will be used to define time intervals.

Overall time on study will be defined as the total number of days a participant is known to be followed on study calculated as follows:

Overall time on Study = (Last date on study) – (Date of first dose) +1

Last date on study is defined as the date of the latest visit or evaluation or telephone contact, or time of death from all available data for a given participant.

Disease duration is defined as time from age at symptom onset to age at informed consent.

3.2. Study Treatments

The main presentations of data will use the following groups:

Baseline characteristics displays for assessing sham and 232SM203 similarity.

CS3B	CS3B	232SM203		
Sham control	Matched Sham control	12/12 mg nusinersen	50/28 mg nusinersen	Total

Baseline characteristics, disposition, exposure, time on study

Matched CS3B	232SM203		
Sham control	12/12 mg nusinersen	50/28 mg nusinersen	Total

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

For summaries of post-baseline efficacy by visit/time

Matched CS3B	232SM203	
Sham control	12/12 mg nusinersen	50/28 mg nusinersen

Primary and secondary endpoints analyses comparing sham and key adverse event displays

Matched CS3B	232SM203
Sham control	50/28 mg nusinersen

For secondary endpoints comparing higher dose to 12/12 mg in 232SM203, safety and post baseline safety parameters:

232SM203	
12/12 mg nusinersen	50/28 mg nusinersen

[redacted]

3.3. Study Periods

The total study duration for each participant will be approximately 323 to 420 days

- Screening: 21 days
- Loading period: 64 days

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

- Maintenance period: 215 days
- Follow-up: 23 to 120 days

3.4. Key Derived Variables

Not applicable.

3.5. Stratification Factors and Subgroup Variables

Stratification Factors

Randomization was stratified as follows:

- For participants with infantile-onset SMA by disease duration:
  - ≤ 12 weeks and
  - > 12 weeks

Subgroup Variables

The following are defined as subgroups:

- Baseline CHOP INTEND (<= median, > median, median calculated 232SM203 ITT Set)
- Geographic region:
  - Asia-Pacific (Australia, China, Hong Kong, Japan, Lebanon, Saudi Arabia, South Korea, Taiwan, Turkey)
  - Europe (Belgium, France, Germany, Great Britain, Hungary, Italy, Poland, Russia, Spain, Sweden)
  - North America (Canada, United States)
  - South/Central America (Brazil, Chile, Colombia, Mexico)

The main analyses for the primary endpoint and time to death and time to death or permanent ventilation for higher dose versus sham will be repeated for the subgroups: disease duration and baseline CHOP INTEND total score. Summaries of change from baseline in CHOP INTEND will be repeated for each of the three subgroups up to Day 183.

The main analyses for the change in CHOP INTEND to Day 302 and time to death and time to death or permanent ventilation will be repeated for subgroups: disease duration, baseline CHOP INTEND for the higher dose versus 12 mg comparison within 232SM203.

Summaries of change from baseline in CHOP INTEND by visit will be presented for each of the three subgroups for the higher dose versus 12 mg comparison within 232SM203.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

3.6. Analysis Sets

3.6.1. Within 232SM203

The Safety Set

The Safety Set is defined as all participants who receive at least one dose of nusinersen; participants will be analyzed in the treatment group based on what they actually received.

The Intent-to-Treat (ITT) Set

The Intent-to-Treat (ITT) Set is defined as all participants who are randomized and receive at least one dose of nusinersen; participants will be analyzed in the treatment group to which they are randomized. This will be the primary population for the analysis of efficacy endpoints.

[redacted]

Per Protocol Set

A Per Protocol Set will include the subset of the ITT Set who have no significant protocol deviations that would be expected to affect efficacy assessments.

Significant protocol deviations will be determined prior to database lock and will include:

Participants who have an SMN copy number other than 2, participants who were greater than 7 months of age at Screening, age of onset of clinical sign >180 days, participants who had hypoxemia at baseline (< 96% at an altitude of < 1500 meters, < 92% at an altitude of 1500 to 2000 meters, or < 90% at an altitude > 2000 meters at Screening).

Participants who are on permanent ventilation or tracheostomy at Screening.

Participants who took experimental or approved agents for the treatment of SMA, including gene therapy, biological agent, or device, during the study. This includes marketed agents at experimental doses that are being tested for the treatment of SMA (oral albuterol/salbutamol, riluzole, carnitine, sodium phenylbutyrate, valproate, hydroxyurea, etc.).

Participants who had a previous diagnosis that would confound the data/conflict with SMA.

If participants have missed planned doses/sham but still completed study.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

3.6.2. External data

For CS3B, the efficacy set at Day 183 and Day 302 as defined for the final analysis of CS3B will be utilized.

The following table describes analysis sets for the historical data and how they will be referenced in the remainder of the document.

Table 1: Analysis set

Analysis Set Label	Description
CS3B Day 183 efficacy set	This is the 'Evaluable set for Day 183' referenced in the Final CS3B CSR. It consists of N=37 sham control and N=73 12 mg nusinersen participants.
CS3B Day 302 efficacy set	This is the 'Evaluable set for Day 302' referenced in the Final CS3B CSR. It consists of N=28 sham control and N=49, 12 mg nusinersen participants.
Matched Sham Set	Analysis set for main comparisons of higher dose versus sham control. Subset of the 37 sham control participants in the CS3B Day 183 efficacy set similar to higher dose 232SM203 participants and all higher dose participants in the ITT set.
Matched Sham Set (sensitivity)	Subset of the 37 sham control participants in the CS3B Day 183 efficacy set similar to all N=75 232SM203 participants in the ITT set.
Greedy matched set (sensitivity)	Set of CS3B sham and 232SM203 higher dose participants matched by greedy matching. Note that not all higher dose participants will have a match.

[redacted]

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

[redacted]

4. List of Planned Study Analyses

4.1. Interim Analysis

No interim analysis is planned. The unblinding plan details the steps to ensure that the blind is maintained throughout the entire study.

4.2. Primary Analysis

The primary analysis is to test the mean change in CHOP INTEND total score from baseline to Day 183 for the 50/28-mg in 232SM203 compared to the CS3B matched sham control Group, using the joint-rank test and handling drop out not due to death with multiple imputation.

4.3. Final Analysis

The final analysis will be performed when all infantile onset participants in Part B have completed the follow-up period.

5. Statistical Methods for Planned Analyses

Within the infantile-onset population to control the overall type 1 error at a 2-sided alpha level of 0.05, a sequential testing procedure ranked in the order of the primary and secondary endpoints will be utilized. Table 2 summarizes the planned statistical testing. Inferential conclusions about each successive analysis require statistical significance of the prior one.

Table 2: Sequential Testing Procedure for Primary and Secondary endpoints

Rank	Endpoint	Comparison Higher dose to:	Population	Description
1	Change from baseline to Day 183 in CHOP-INTEND total score	CS3B Sham control	Matched Sham Set	Primary
2	Proportion of HINE Section 2 motor milestone responders at Day 183	CS3B Sham control	Matched Sham Set	Secondary
3	Change from baseline to Day 183 in HINE Section 2 total score	CS3B Sham control	Matched Sham Set	Secondary
4	Change from baseline to Day 183 in plasma NF-L	CS3B Sham control	Matched Sham Set	Secondary

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

Rank	Endpoint	Comparison Higher dose to:	Population	Description
5	Change from baseline to Day 302 in CHOP-INTEND total score	232SM203 12 mg	ITT set	Secondary
6	Change from baseline to Day 302 in HINE Section 2 total score	232SM203 12 mg	ITT set	Secondary
7	Change from baseline to Day 64 in plasma NF-L	232SM203 12 mg	ITT set	Secondary
8	Time to death or permanent ventilation	CS3B Sham control	Matched Sham Set	Secondary
9	Time to death	CS3B Sham control	Matched Sham Set	Secondary
10	Time to death or permanent ventilation	232SM203 12 mg	ITT set	Secondary
11	Time to death	232SM203 12 mg	ITT set	Secondary

5.1. General Principles

Descriptive summary statistics will be presented for all primary, secondary [redacted] endpoints collected. Unless otherwise specified, for continuous endpoints, the summary statistics will generally include number of participants with data, mean, standard deviation, median, 25th percentile, 75th percentile, minimum and maximum. For categorical endpoints, the summary statistics will generally include number of participants with data, and the percentage of those with data in each category.

All summaries and listings will be presented by study groups unless otherwise specified. Visits in listings will be displayed as per CRF data collection rather than analysis visits.

The statistical software, SAS®(Version 9.4) will be used for all summaries and statistical analyses.

Visit Windows for Early Withdrawal, Unscheduled Visits

Data from early withdrawal visits, post baseline unscheduled assessments will be assigned to an appropriate scheduled post baseline visit using a windowing scheme for assessments that are tabulated or summarized by visit. Scheduled visits will not be windowed.

The visit windowing will be performed for efficacy and safety endpoints.

For safety endpoints the following windows will be utilized.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

Table 3: Visit windows for safety Endpoints in 232SM203.

Visit	Lower Bound	Upper Bound	Target Day
Baseline			<=2
Day 15	2	21	15
Day 29	22	47	29
Day 64	48	99	64
Day 135	100	159	135
Day 183	160	232	183
Day 279	233	290	279
Day 302	291	348	302

For efficacy endpoints, if the actual study day falls within: +/- 7 days of scheduled post baseline loading dose visits and +/-14 days of a maintenance dose then it will be considered for windowing.

The following rules will be implemented for windowing of these visits for both safety and efficacy.

- If more than one observation is within the same window, data from the regular scheduled visit will be used for that visit.
- If neither of the observations are from a regular scheduled visit the observation closest to the planned target date will be used. However, if both the observations are equidistant from the target date, the latest one will be used; if repeated measurements are on the same day, then the last measurement will be used.
- If windowing safety data and there is more than one observation in a window for a dosing visit the observation on the day of dosing will be chosen over the observation where dosing did not occur on the same day.
- In windowing laboratory data, if a central and local result is available within the window then the central laboratory result will be chosen.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

5.2. Participant Accountability

The number (and percentage) of subjects screened, screening failure, randomized, withdrew prior to dosing, dosed, completed treatment, discontinued treatment and the reasons for discontinuation from treatment, completed the study, withdrew the study and reason for withdrawal will be summarized in a table.

If there are any subjects who discontinued from treatment or withdrew from study due to reasons associated with the COVID-19 pandemic or reasons associated with humanitarian emergencies, a separate summary will be presented to summarize these reasons.

A listing of subjects who discontinued treatment/withdrew from study and the associated reasons for discontinuation/withdrawal will be presented. Two separate listings will also be presented for subjects who discontinued treatment/withdrew from study due to the COVID-19 pandemic or to humanitarian emergencies if applicable.

5.3. Demographic and Baseline Characteristics

Baseline data (demography, medical history, SMA history, and baseline disease characteristics) will be summarized.

Formal statistical analyses will not be done to test for homogeneity between treatment groups. If there are apparent heterogeneities between the groups in any of the participant characteristics that are of clinical importance or could affect the treatment outcome, the impact of the imbalances will be investigated and, if appropriate, adjustments made in the efficacy and safety analyses.

Demography includes age at screening, age at first dose, sex, ethnicity, race and country,

Medical history will be coded in MedDRA and the number and percentage of subjects with each medical history presented by preferred term.

SMA history will include gestational age, birth weight, age at symptom onset, time from disease onset to enrollment, age at SMA diagnosis, history of SMA symptoms (hypotonia, developmental motor delay, paradoxical breathing, pneumonia or respiratory symptoms, limb weakness, swallowing or feeding difficulties, and other symptoms), and the number of copies of the SMN1 and SMN2 gene.

Baseline disease characteristics will be assessed by CHOP INTEND total score, total HINE Section 2 motor milestones, and growth parameters - these assessments are further described in Section 6. If a participant was receiving ventilation at Day 1, then the number of hours of ventilation use on Day 1 will also be summarized as baseline. In addition, use of a gastrointestinal tube will be presented. Use of gastrointestinal tube will be determined from either medical history or as a procedure prior to first dose coded as ‘Gastrostomy’ or ‘Gastrointestinal tube insertion’.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

Demographic and baseline disease characteristics will be presented for the ITT Set, the PPS, the Safety Set as appropriate. Additional presentations will be presented to facilitate comparisons at baseline for the planned efficacy comparisons. For example, for the primary endpoint to compare the higher dose to CS3B sham control.

5.4. Protocol Deviations

Protocol deviations identified during site monitoring will be captured in a protocol deviation log and categorized as major or minor deviations. These deviations will be listed. Major protocol deviations will be summarized including those that lead to exclusion from the Per protocol Set.

Participants with incorrect stratification i.e. any mismatch between the stratification used for randomization on the IRT system and the actual stratification according to the eCRF will also be listed. A separate summary of major protocol deviations related to the COVID-19 pandemic will be presented.

5.5. Study Treatment Exposure

The number of doses received, and the number of sham procedures performed will be displayed using frequency distributions. The cumulative amount of nusinersen received will be summarized using summary statistics.

The overall time on study will be summarized descriptively by treatment. Overall time on study will be calculated (section 3.1). The duration will also be categorized and summarized using the following categories: <29 days, >=29 to 64 days, >=64 to 99 days, >=99 to 135 days, >=135 to 183 days, >=183 to 218 days, >=218 to <260, >=260 to 302 and >=302 days

Given the long half-life of nusinersen, participants are considered to be exposed from the time the first dose of nusinersen was administered (in or before the start of study) to the date of last visit or contact.

Separate listings will be provided showing what subjects were randomized to as well as study drug administration data which will include lot numbers, actual treatment received, cumulative number of doses and cumulative doses. A listing of dosing errors will also be provided i.e. where either the incorrect dose was administered as specified in protocol deviation log and kit description from IRT system. A listing will also be presented to show whether any of the missed doses are related to COVID-19. If any of the doses are due to COVID-19, then the number of impacted doses will be presented.

6. Efficacy Endpoints

6.1. General Analysis Methods for Efficacy Endpoints

Multiple imputation (MI)

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

The multiple imputation method [Schafer 1997, Schafer 1999] will be used for participants who discontinue for a reason other than death or have a missing efficacy result at the timepoint of interest. This will be performed for CHOP INTEND, HINE Section 2 total score and NF-L. Any imputation on an item level within a scale for CHOP INTEND and HINE section 2 will be performed prior to performing MI.

The imputations will be performed on total scores. All available data will be used i.e., where subjects have missed doses but continued to perform assessments; otherwise, any missing values will be imputed using MI. In a situation where historical data is used as the control group only the visits which are common between the two groups will be utilized. For example, for the CS3B sham control versus higher dose comparison only the Day 64, Day 183 and Day 302 visits will be included in the list of post baseline visits.

In this situation if convergence issues occur then only the Day 64 and Day 183 post baseline visits will be included.

The Markov Chain Monte Carlo (MCMC) method will be used to impute the missing score.

The treatment variable will be coded so that the Higher dose group is the first in the sort order. Prior to the Proc MI step the dataset will be sorted by treatment and then ascending unique subject identifier (USUBJID). The variable list in the model for imputations will include:

- treatment group, disease duration, baseline value for the endpoint, and all available post baseline values
- For NF-L, the log of baseline and post-baseline values will be used in the PROC MI step.

A set of 100 complete imputed datasets will be generated and the relative efficiency parameter will also be checked to determine the acceptability of the imputed results i.e. it should be close to 98% or higher. The seed used for the multiple imputation for CHOP INTEND will be 5846892, for HINE section 2 Motor Milestones will be 8746890 and for NF-L will be 7635789.

If the MI procedure imputes any values outside the expected range for the scale, e.g. from 0 to 64 for CHOP INTEND, then values below 0 will be set to 0 and values above 64 will be set to be 64. For NF-L if any imputed values are < log(LLOQ) then set as log(LLOQ/2).

For each of the 100 imputed datasets, the endpoint will be compared between treatment groups using an ANCOVA model for continuous endpoints and logistic regression for binary endpoints as described in the main analysis for each endpoint. Additional detail such as how the joint rank test will be implemented and the adjustment for baseline covariates is described below and for each endpoint in Section 6.3 and 6.4

The estimates from the 100 fitted models will be combined to provide an overall estimate with corresponding confidence intervals and p-value using PROC MIANALYZE [Little et al, 2002].

For the continuous endpoints, the difference between treatments and the corresponding CI will be presented. For the analyses using logistic regression the odds ratio and corresponding CI will

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

be presented. In addition, the estimates of the binomial proportions in each treatment arm and the differences between treatments will be presented (Ratitch, 2013).

In the presentation of results from multiple imputed data, the number of subjects with missing data will be summarized.

Joint rank methodology

Death is a potential outcome in the infantile-onset population and the joint rank methodology will be utilized to incorporate both the functional endpoint and mortality.

The joint rank methodology [Berry 2013] allows for a statistical test of the treatment effect on the endpoint while accounting for truncation of data due to deaths. In this analysis, a subject's joint rank score will be calculated by comparing each subject to every other subject in the study, resulting in a score of +1 if the outcome was better than the subject being compared, -1 if worse, and 0 if the same. The subject's score will then be calculated by summing their comparison to all the other subjects in the study.

For the purpose of this calculation, considering an endpoint for change to Day 183, subjects will be grouped into the following 2 categories:

- Group 1: Participants who complete Day 183 and have data available AND subjects who withdraw from the study due to reasons other than death and don't have Day 183 data available
- Group 2: Participants who die will be ranked based on the time of death

The criteria for withdrawals are based on withdrawal from study. A subject may discontinue treatment but still continue in the study and in this case data from assessments conducted after treatment will be included for analysis.

Joint rank methodology + MI

MI will be used to impute all missing data including data after withdrawal from the study or where a participant completed the study but doesn't have an evaluable score at the timepoint of interest. No efficacy data would be available after death. MI will be performed as described earlier and subjects who died will be included in the MI model, but any values imputed after death will not be used for determining the rank score. Therefore, all subjects except withdrawals due to death will have a value at Day 183 (either observed or imputed). In each of the 100 imputed datasets, subjects will be ranked as follows:

- Subjects in Group 2 will be given lower ranks than subjects in Group 1, with the lowest ranks being given to the subjects who die in the shortest time after first dose. Progressively higher ranks will be given to subjects who die at longer times after first dose.
- Subjects in Group 1 will rank higher than subjects in Group 2. Progressively higher ranks will be given to subjects with a higher change in CHOP INTEND at Day 183.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

For example, for 2 subjects in Group 1 their comparison score will be based on the endpoint of Day 183 CHOP INTEND. The subject being ranked will be given a comparison score of -1 if their CHOP INTEND is worse than that of the comparison subject. If the score is the same in both subjects their comparison will be considered a tie and so a comparison score of 0 will be assigned. For subjects who do not survive, a subject who dies earlier than the comparator subject will be given a comparison score of -1. If two subjects die the same number of days from their first dose of study medication their comparison will be considered a tie i.e. comparison score of 0 will be assigned.

The subjects will be ranked according to their subject rank scores in each of the 100 MI datasets. In the instance of ties, the mean of the corresponding ranks will be assigned to the tied subjects. Hence, in general, these comparisons will result in subjects who die being assigned the worst scores and ranked according to time of death. Subjects who survive and complete the study and subjects who withdraw from the study will be ranked more favorably than subjects who die.

For each of the 100 imputed datasets, the ranked scores will be compared between treatment groups using an ANCOVA model. The ANCOVA model will include treatment as a fixed effect and adjusted for each of the corresponding baseline value for the endpoint and disease duration [redacted]

The model will be used to obtain the p-value from PROC MIANALYZE for the joint rank.

Adjustment of ranking criteria for JRT for withdrawals not due to death

In order to test the sensitivity of handling withdrawals not due to death, ranking criteria will be adjusted as a sensitivity analysis for JRT. Instead of 2 categories, subjects will be grouped into 3 categories:

- Group 1: Participants who complete Day 183 and have data available or complete the study but don't have an evaluable result for Day 183
- Group 2: Participants who withdraw from the study due to reasons other than death and don't have Day 183 data available
- Group 3: Participants who die will be ranked based on the time of death

Subjects will be ranked as follows:

- Participants in Group 3 will be given lower ranks than subjects in Groups 1 and 2, with the lowest ranks being given to the subjects who die in the shortest time after first dose. Progressively higher ranks will be given to subjects who die at longer times after first dose.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

- Participants in Group 2 will rank lower than subjects in Group 1, with the lowest ranks being given to the subjects who withdraw from the study in the shortest time after first dose. Progressively higher ranks will be given to subjects who withdraw from the study at longer times after first dose.
- Subjects in Group 1 will rank higher than subjects in Groups 2 and 3. Progressively higher ranks will be given to subjects with a higher change from baseline at Day 183 i.e. smaller decline at Day 183.

The same approach using the joint rank test with MI will be implemented for the analysis. The main difference is that the imputed values will only be utilized for ranking if the participant is in group 1. Otherwise, the ranking will be performed based on time in study. [redacted]

Adjustment of ranking criteria for JRT for permanent ventilation

In order to examine the impact of considering permanent ventilation in evaluation of this endpoint the ranking criteria will be adjusted as sensitivity analysis for JRT. Subjects will be grouped into 2 categories:

- Group 1: Participants who complete Day 183 and have data available or reach beyond Day 183 and don't have an evaluable result for Day 183
- Group 2: Participants who die or are confirmed to have met permanent ventilation, i.e. tracheostomy or 22nd day of ventilation is on or prior to Day 183. Participants will be ranked based on the earlier of time of death or time of meeting permanent ventilation definition.

The same approach using the joint rank test with MI will be implemented for the analysis. The main difference is that the imputed values will only be utilized for ranking if the participant is in group 1. Otherwise, the ranking will be performed based on time in study. [redacted]

Rank based implementation of Joint rank method + MI

As a Sensitivity analysis, a new version of the joint rank test will be implemented, using the change in ranks rather than directly the survival time and change in CHOP INTEND.

Initially, all participants will be ranked based on the baseline CHOP INTEND score, so for the 12 mg versus 50/28mg comparison within 232SM203, with a total of N=75, the lowest score will be ranked as 1 and the highest as 75.

Group 1 will be defined as participants who died prior to Day 183. For this group, ranking will be performed based on Study Day of Death.

Group 2 will be defined as participants who did not die and have either an imputed or observed Day 183 result for total CHOP INTEND.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

Subjects will be ranked in each of the 100 MI datasets as follows. The subjects will be ranked for Day 183 with subjects who died in group 1 given a lower rank than subjects in group 2. The subject who died first will be given a rank of 1. For group 2 the lowest Day 183 CHOP INTEND score will be given the lowest rank and the highest CHOP INTEND score will be assigned the highest rank. In the case of ties the mean of the ranks will be assigned to tied subjects

For each of the 100 imputed datasets the change in ranks will be compared between treatment groups using an ANCOVA model. The response variable will be determined as the difference: Rank at Day 183 – [redacted]. The model will include treatment as a fixed effect and adjusted for each of the corresponding baseline value for the endpoint and, disease duration. [redacted]

ANCOVA model

For each of the 100 imputed datasets, the endpoint will be compared between treatment groups using an ANCOVA model for continuous endpoints (primary, secondary and NF-L endpoints). The ANCOVA model will include covariates for each of the corresponding baseline value for the endpoint, baseline CHOP INTEND total score (if not the endpoint under consideration) and disease duration at informed consent. The model will be used to present overall LS means and standard errors for each treatment group, and LS mean differences for treatment with corresponding 95% CI.

ANCOVA model without equality of variances assumption

As a sensitivity analysis the main analysis will be repeated using the repeated statement in PROC Mixed to estimate the variance for each treatment group separately. For this analysis the degrees of freedom will be adjusted using the Satterthwaite option.

Jump to reference analysis for MI with the Joint Rank test

Jump to reference (J2R) is a sensitivity analysis such that imputed values for patients in the higher dose arm take on attributes of the reference arm (matched sham subjects) after discontinuation. This is a conservative assumption for the higher dose versus sham comparisons as due to the long half-life it is not likely that following withdrawal the benefit of treatment will completely disappear. J2R analysis will be implemented using the SAS macro RMCONJPLUS (https://www.lshtm.ac.uk/research/centres-projects-groups/missing-data#dia-missing-data).

In the joint rank step as for the main analysis, for patients who died the ranking will be performed based on day of death and then for patients who have not died the ranking will be based on the change from baseline in efficacy score (imputed or observed) at the time point under consideration.

Trimmed mean

As a supplemental analysis, the change from baseline in CHOP INTEND total score between treatment groups will be compared using the trimmed mean method to account for informative censoring due to death or withdrawal [Permutt and Li 2015].

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

The trimmed mean is an averaging method which eliminates a partial percentage of the smallest values before evaluating the standard mean of the given data. The trimmed mean difference can be interpreted as the mean difference of the top fraction of responses between the two groups. To do this, responses are ranked and subjects who have died or withdrew are considered as having worse outcomes than those who survived and completed the trial. To ensure that subjects who died or withdrew receive the lowest rank in response an arbitrary value of -999 will be assigned. The choice of this value does not impact the analysis results because these subjects are eventually excluded from the calculation of the trimmed mean statistics. Subject response scores in the two groups will then be ranked from the largest to the smallest separately. The top 100*p% response scores from the two groups will be retained and the mean response difference is calculated as the trimmed mean statistics. Here, p is the minimum of the proportions of subjects who survived and completed the assessment in the two groups.

As a hypothetical example, using the 12 mg versus 50/28 mg with 232SM203 comparison, assume at the end of the study in which 75 subjects were randomized and dosed, the number of subjects who died or withdrew is 4 out of 50 (8.0%) in the nusinersen 50/28 mg group and 5 subjects out of 25 (20%) in the control group. The 9 subjects who died or withdrew in the two groups are assigned a response score of -999. The CHOP INTEND scores at Day 183 in the two groups are then ranked from largest to smallest separately. The top 80% of scores from each group will be retained and the mean difference is then calculated as the trimmed mean statistic.

Note that MI will not be used for trimmed mean approach. For subjects who don't have Day 183 data available (who do not withdraw from study or die), they will also be trimmed from analysis, i.e. they will also be assigned the arbitrary value of -999.

The p-value associated with the observed test statistic is then calculated using a reference distribution generated by the re-randomization procedure described below:

1. For all 75 subjects randomized in the ITT population, randomly assign subjects in a 50/25 ratio to treatment (nusinersen 58/28 mg or control) such that there are 50 subjects in the nusinersen 58/28 mg group and 25 in the control group. This can be done using simple random sampling without replacement. The seed used will be 5846894.
2. Calculate the trimmed-mean test statistic as described above using the random treatment assignment in Step 1
3. Repeat Steps 1 and 2 N times (N=10,000). As a result, N trimmed-mean statistics will be generated. The empirical distribution of the N trimmed-mean statistics will then serve as the reference distribution for p-value calculations.

Denote the observed trimmed mean statistic as T0. Once the reference distribution has been established, the p-value associated with the observed test statistic is calculated as proportion of simulated test statistics that are either greater than |T0| or less than - |T0|.

Responder Analyses of CHOP INTEND and HINE Section 2

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

For CHOP INTEND it is planned to use multiple imputation to deal with missing data and the analysis will be performed using logistic regression. Should the number of responders be less than 5 in either group, Fisher's exact test will be used instead. In this scenario, it will not be possible to use multiple imputation and the following approach will be implemented to impute a missing response:

For a given timepoint such as Day 183, subjects who die or withdraw from the study prior to Day 183 will be counted as non-responders and will be included in the denominator for the calculation of the proportion of responders. In determining the response at Day 183 where a subject has missing response but the two immediate flanking visits both have an evaluable response an imputation will be performed. For example, the visit of interest is Day 183 and the immediate flanking visits are Day 135 and Day 279. If at both Day 135 and Day 279 the subject is a responder, then an imputation will be performed to count the missing Day 183 as a responder. For any other combination of missing/non response observed in flanking visits the subject will be considered a non-responder.

For HINE Section 2 motor milestone responders it is not planned to use multiple imputation. Therefore, if the endpoint is at Day 302 and the participants has a missing HINE Section 2 response, they will be considered to be a non-responder. In a situation where the endpoint is at Day 183 and flanking visits may be available the imputation approach described above for CHOP INTEND response will be implemented.

6.2. Use of Historical Data

The primary and some secondary endpoints will compare high dose to sham control. A total of N=37 are available from CS3B at Day 183. The following algorithm will be used to select the maximum number of Sham control subjects who are comparable at baseline to 232SM203 (N=75) based on two covariates: disease duration at informed consent, baseline CHOP INTEND total score. The algorithm will stop at a minimum of 20 participants.

[redacted]

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

[redacted]

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

[redacted]

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

[redacted]

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

[redacted]

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

[redacted]

6.3. Primary Efficacy Endpoint

6.3.1. CHOP INTEND

The CHOP INTEND infant motor function scale is comprised of 16 test items, nine of which are scored 0, 1, 2, 3, or 4 with greater scores indicating greater muscle strength, five are scored as 0, 2, or 4, one is scored as 0, 1, 2, or 4, and one as 0, 2, 3, or 4. This can result in a worst possible total score of 0 to a best possible total score of 64. CHOP INTEND is used to assess spontaneous movement in the upper extremities, spontaneous movement in the lower extremities, hand grip, head in midline with visual stimulation, hip adductors, rolling elicited from the legs, rolling elicited from the arms, shoulder and elbow flexion and horizontal abduction, shoulder flexion and elbow flexion, knee extension, hip flexion and foot dorsiflexion, head control, elbow flexion, neck flexion, head/neck extension, and spinal incurvation.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

The baseline for CHOP INTEND is defined as the mean of the total score assessments taken during the screening/baseline period prior to the first dose.

Missing Values

Missing data will be imputed on an individual CHOP INTEND test item level to provide a total score only if at least one item is assessed.

If a test item is missing at screening, then the missing value will be imputed as the median of the non-missing values of the stratum to which the subject belongs to: age at symptom onset (≤12 weeks, > 12 weeks) by disease duration (≤12 weeks, >12 weeks). Specifically, the four strata are:

- Age at symptom onset ≤12 weeks and disease duration ≤12 weeks
- Age at symptom onset ≤12 weeks and disease duration >12 weeks
- Age at symptom onset >12 weeks and disease duration ≤12 weeks
- Age at symptom onset >12 weeks and disease duration >12 weeks

If, for the subject with missing CHOP INTEND test items at a particular visit, the corresponding visit is flanked by visits with non-missing test items, then the missing value for those test items will be imputed using linear interpolation with the result rounded to the nearest integer score. Otherwise, if the missing visit is the last visit, missing test items will be imputed as the lowest value in the stratum (age at symptom onset by disease duration) to which the subject belongs within the same treatment group at the same visit.

In the event of no observed data for imputation, disease duration by treatment will be used as the classification factor for the purpose of identifying non-missing data for imputation.

Of note, only observed data will be utilized for imputation purposes. Missing CHOP INTEND items will be imputed first prior to any analysis.

Main Analysis

The primary efficacy endpoint of change from baseline to Day 183 in CHOP INTEND total score will be analyzed using the joint-rank methodology to account for mortality [Berry 2013]. This joint-ranking procedure allows for a statistical test of the treatment effect on the CHOP INTEND total score while accounting for loss of data due to deaths. Multiple imputation (MI) will be used to handle withdrawals, when implementing the JRT methodology. Any imputation on an item level within a scale will be performed prior to performing MI. The estimates will be obtained from an ANCOVA model for change from baseline in CHOP INTEND total at Day 183 with missing data imputed using multiple imputation. Note that all available data up to Day 302 will be included in MI procedure.

The main analysis of the primary endpoint of Day 183 change from baseline in CHOP INTEND is the composite estimand [ICH E9 (R1) Addendum 2017].

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

The estimand of the primary analysis is defined as:

- Population: all Higher Dose subjects in the ITT population and CS3B sham Day 183 matched efficacy set
- Variable: change from baseline to Day 183 in CHOP INTEND total score OR Study day of death
- Intercurrent events: handled using a composite strategy in which subjects who have intercurrent events are ranked against each other and against subjects without any intercurrent event using the joint rank methodology based on MI datasets as detailed in Section 6.1. The relevant intercurrent events are death and withdrawals. For deaths, ranking will be based on time on study and for withdrawals not due to death ranking will be based on the imputed Day 183 value from the MI datasets.

The ANCOVA model will include covariates for baseline CHOP INTEND total score and disease duration at screening. [redacted] The P-value for the difference between treatment groups based on the rank scores at Day 183 will be calculated from PROC MIANALYZE.

Sensitivity analyses

- Repeated using CS3B sham Day 183 matched to 232SM203 efficacy set (sensitivity)
- Repeated using Greedy matched set (sensitivity)
- Adjustment of ranking criteria for JRT for withdrawals not due to death

  In order to test the sensitivity of handling withdrawals not due to death, ranking criteria will be adjusted as a sensitivity analysis for JRT
  These analyses will be performed as described in Section 6.1
- Adjustment of ranking criteria for JRT for permanent ventilation
- Main analysis will be repeated using an ANCOVA model where the variance in each treatment group is estimated separately

In order to examine the impact of considering permanent ventilation in evaluation of this endpoint the ranking criteria will be adjusted as sensitivity analysis for JRT. These analyses will be performed as described in Section 6.1.

- Exclusion of deaths from MI model

  The MI model will be run for CHOP INTEND total score for the ITT population excluding data from any subjects who died prior to Day 183, so that missing data for survivors are not influenced by any systematic difference there may be in data from subjects who died. The joint rank analysis will then be conducted using these imputed datasets for survivors and deaths will be ranked the lowest based on time of death. The p-value from the joint rank analysis with MI will be presented as per the main analysis.

- Data truncated to include only up to and including Day 183

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

In the event that a participant has CHOP INTEND missing items at Day 183 then data from later visits is utilized as part of interpolation. Some of the subjects in the CS3B study only have data up to Day 183 before the study was stopped in contrast to 232SM203 where all subjects potentially could be followed to Day 302. In order to investigate the sensitivity of the analysis to this truncation an additional analysis will be performed where only the data up to and including Day 183 is utilized from both CS3B and 232SM203 and the item level imputation approach described in Section 6.1 is applied.

- The main analysis will be repeated incorporating Jump to reference in the MI step.
- The main analysis will be repeated for the following subgroups:
  - disease duration
  - baseline CHOP INTEND

Supplementary Analyses

In order to provide additional understanding of the treatment effect, two supplementary analyses will be performed on the primary endpoint. The first, in order to provide an estimate of the difference in CHOP INTEND using an ANCOVA model at each timepoint the MI dataset generated for the primary endpoint will be utilized. As a sensitivity analysis, this will be repeated using an ANCOVA model where the variance in each treatment group is estimated separately.

The second will utilize the following estimand which is described as the trimmed mean:

- Population: The ITT Set is considered and the treatment arm with the highest % of participants withdrawn by Day 183 is determined (labelled X). The same fraction is applied to the other treatment arm and the analysis is performed on (100-X) % of the ITT Set
- Variable: difference between treatments in the endpoint means in the (100-X) % subset of participants with the most favorable outcomes,
- Handling of intercurrent events: Due to the population definition, the intercurrent events of death or withdrawal from study are dynamically trimmed from the ITT Set. The assumption here is that any ICE is an equally bad outcome.
- Population level summary – ANCOVA

These analyses are further described in Section in Section 6.1.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

Table 4: CHOP INTEND: Change from baseline to Day 183, accounting for mortality/dropout. Comparing the 50/28-mg Group to Study CS3B Sham Control

Type	Analysis Method	Analysis Population: Further Notes
Primary Analysis	Joint rank using multiple imputation (MI)	Matched Sham Set
Sensitivity Analysis	Joint rank using multiple imputation (MI)	Matched Sham Set (sensitivity)
	Joint rank using multiple imputation (MI)	Greedy matched set (sensitivity)
	Adjustment of ranking criteria for JRT for withdrawals not due to death	Matched Sham Set
	Adjustment of ranking criteria for JRT for permanent ventilation	Matched Sham Set
	Subjects who died will be excluded from the MI model	Matched Sham Set
	Joint rank using multiple imputation (MI)	Matched Sham Set. Data truncated to only include data up to Day 183 (inclusive).
	Joint rank using Jump to Control MI	Matched Sham Set
	Rank based implementation of joint rank method + MI	Matched Sham Set
	Joint rank test using multiple imputation (MI), ANCOVA model not assuming equality of variance	Matched Sham Set
	Joint rank using multiple imputation (MI) by disease duration	Matched Sham Set
	Joint rank using multiple imputation (MI) by baseline CHOP INTEND	Matched Sham Set
Supplementary Analysis	Trimmed mean	Matched Sham Set

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

Endpoint and Comparison	Analysis Method	Analysis Population: further notes
Supplemental analysis		
Difference (50/28mg- Matched Sham Control) in change in CHOP INTEND from baseline by visit	ANCOVA using MI	Matched Sham Set
Difference (50/28mg- Matched Sham Control) in change in CHOP INTEND from baseline by visit	ANCOVA using MI (Not assuming equality of variances)	Matched Sham Set

6.4. Secondary Efficacy Endpoints

6.4.1. Proportion of motor milestone responders at Day 183: comparison: 50/28mg versus CS3B sham control

Hammersmith Infant Neurological Examination (HINE) section 2

Motor milestones will be determined using Section 2 of the Hammersmith Infant Neurological Examination (HINE) which is comprised of eight motor milestone categories: voluntary grasp, ability to kick in supine position, head control, rolling, sitting, crawling, standing and walking.

Table 5: Hammersmith Infant Neurological Examination Section 2 - Motor Milestones

Motor Milestone Category	Milestone Level Progression Score (Age Expected in Heathy Infants1) → Improvement

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

	0	1	2	3	4
Voluntary grasp	No grasp	Uses whole hand	Finger and thumb/ immature grasp	Pincer grasp	
Ability to kick (in supine)	No kicking	Kicks horizontal; legs do not lift	Upward (vertically) [3 months]	Touches leg (4 to 5 months)	Touches toes (5 to 6 months)
Head control	Unable to maintain upright (< 3 months)	Wobbles (4 months)	All the time upright (5 months)		
Rolling	No rolling	Rolls to side (4 months)	Prone to supine (6 months)	Supine to prone (7 months)	
Sitting	Cannot sit	Sits with support at the hips (4 months)	Props (6 months)	Stable sit (7 months)	Pivots (rotates) [10 months]
Crawling	Does not lift head	On elbow (3 months)	On outstretched hand (4 to 5 months)	Crawls flat on the abdomen (8 months)	On hands and knees (10 months)
Standing	Does not support weight	Supports weight (4 to 5 months)	Stands with support (8 months)	Stands unaided (12 months)	
Walking	No walking	Bouncing (6 months)	Cruising (holding	Walks independently (15 months)	

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

on)
[11 months]

Values for healthy infants in [De Sanctis 2016; Haataja 1999]

Within each motor milestone category (depicted as Motor Milestones in Table 5), there are 3 to 5 levels that can be achieved (depicted as Milestones Progression, or rows, in the Table above). All 8 Motor Milestones are tested during each assessment. A subject whose results after testing all appear in the first column (No grasp, No kicking, Unable to maintain head upright, etc.) has not achieved any motor milestone. Motor milestone achievement is depicted by movement from the left side of the Table above to right side of the Table, as denoted by the Milestone Progression arrow in the Table.

Missing Values

Missing data will be imputed on an individual milestone level only if at least one milestone is assessed at the visit.

If a motor milestone is missing at baseline, then the missing value will be imputed as the median of the non-missing values of the stratum to which the subjects belong to the four strata as defined above CHOP INTEND will be used: age at symptom onset (≤12 weeks, >12 weeks) by disease duration (≤12 weeks, >12 weeks). The median value will be rounded up to the nearest integer value.

If, for the subject with missing motor milestones at a particular visit, the corresponding visit is flanked by visits with non-missing Motor Milestones, the missing value for those motor milestones will be imputed using linear interpolation with the result rounded to the nearest integer score. Otherwise, if the missing visit is the last visit, missing motor milestone value will be imputed as the lowest value in the stratum (age at symptom onset by disease duration) to which the subject belongs within the same treatment group.

In the event of no observed data for imputation, disease duration by treatment will be used as the classification factor for the purpose of identifying non-missing data for imputation.

Of note, only observed data will be utilized for imputation purpose. Missing motor milestone items will be imputed first prior to any analysis.

The definition of a motor milestones responder is based on the motor milestones categories in Section 2 of the Hammersmith Infant Neurological Exam with the exclusion of voluntary grasp using assessment at Day 183 study visit as follows:

(a) Subject demonstrates at least a 2-point increase in the motor milestones category of ability to kick or maximal score on that category (touching toes) whilst having at baseline demonstrated touched leg, or a 1-point

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

increase in the motor milestones category of head control, rolling, sitting, crawling, standing, or walking, AND

(b) Among the motor milestones categories with the exclusion of voluntary grasp, there are more categories where there is improvement as defined in (i) than worsening. Note: for the category of ability to kick, similar to the definition of improvement in (i) above, worsening is defined as at least a 2-point decrease or decrease to the lowest possible score of no kicking. For the other 6 categories, worsening is defined as at lease 1-point decrease.

Main Analysis

- The difference in the proportion of responders between treatment groups at Day 183 will be compared using logistic regression with the total number of motor milestones at baseline, baseline CHOP INTEND and disease duration at screening as covariates. Should the number of responders be less than 5 in either group, Fisher's exact test will be used instead. If Fisher's exact test is used, the unconditional confidence interval for the difference in response rates will be provided [Chan and Zhang 1999].

Sensitivity Analyses

- Subjects who have neither died nor withdrawn with no assessment at Day 183 will be excluded from the analysis.
- An alternative definition of a Motor Milestones responder with the exclusion of voluntary grasp will be used:
  (a) subject demonstrates at least a 2-point increase in the Motor Milestones category of ability to kick or maximal score on that category (touching toes) whilst having at baseline demonstrated touched leg, or a 1- point increase in the Motor Milestones category of head control, rolling, sitting, crawling, standing, or walking, and
  (b) among the motor milestone categories with the exclusion of voluntary grasp, the changes from baseline in total Motor Milestones score is at least 1-point. The total Motor Milestones score is calculated as the sum of scores across motor milestone categories, where 0 is lowest possible score within each Motor Milestones category.
- A Motor Milestones responder is defined as a 2-point increase in the total Motor Milestones score, where voluntary grasp is excluded. For example, assuming no changes in other Motor Milestones categories
  (a) 1-point increase in head control and a 1-point increase in rolling is considered a response,
  (b) a 2-point increase in rolling is considered a response,
  (c) a 2-point increase in rolling with a 1-point decrease in head control is considered a non-response.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

- The main analysis will be conducted excluding subjects with any missing baseline Motor Milestones assessment.

For the analysis at each visit, responder classifications will be based on the motor milestones assessment at the visit of interest. Subjects who die or withdraw prior to the visit of interest will be classified as non-responders. A subject is classified as having died or withdrawn prior to the visit of interest if the last visit prior to death or withdrawal is earlier than the visit of interest.

Table 6: Proportion of Motor Milestones responders at Day 183: comparison: 50/28mg versus CS3B sham control

Type	Analysis Method	Analysis Population: Further Notes
Main analysis	Logistic regression or Fisher's Exact Test	Matched Sham Set
Sensitivity analysis	Logistic regression or Fisher's Exact Test	Matched Sham Set Subjects who have neither died nor withdrawn with no assessment at Day 183 will be excluded.
	Logistic regression or Fisher's Exact Test	Matched Sham Set Net Positive improvement
	Logistic regression or Fisher's Exact Test	Matched Sham Set responder defined as 2-point increase in total milestones score excluding voluntary grasp
	Logistic regression or Fisher's Exact Test	Matched Sham Set (excluding subjects with missing baseline Motor Milestones assessment)

6.4.2. Change in Motor Milestones total score from baseline to Day 183: comparison: 50/28mg versus CS3B sham control.

A total Motor Milestones score for HINE Section 2 will be calculated as the sum of each level and will range from 0 to a maximum score of 26.

Main Analysis

- The change from baseline to Day 183 in Motor Milestones total score will be analyzed using the JRT methodology to account for mortality. The ANCOVA model will include covariates for baseline Total Milestone score, baseline CHOP INTEND

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

total score and disease duration at screening. [redacted]

Sensitivity Analyses

- In order to test sensitivity of data handling for withdrawals not due to death, ranking criteria will be adjusted in the JRT for withdrawals not due to death.
- In order to test if permanent ventilation should be considered in evaluation of endpoint, ranking criteria will be adjusted for JRT to consider permanent ventilation in evaluation.
- The MI model will be run for total score for the ITT population excluding data from any subjects who died.
- Rank based implementation of joint rank method + MI
- The main analysis will be repeated incorporating Jump to reference in the MI step.
- Main analysis will be repeated using an ANCOVA model where the variance in each treatment group is estimated separately

Supplementary analyses

- ANCOVA
- This will be repeated using an ANCOVA model where the variance in each treatment group is estimated separately.

For more details refer to section 6.3.

Table 7: Change from baseline to Day 183 in Motor Milestones total score: comparison: 50/28mg versus CS3B sham control

Type	Analysis Method	Analysis Population: Further Notes
Main analysis	Joint rank using multiple imputation (MI)	Matched Sham Set
Sensitivity analysis	Adjustment of ranking criteria for JRT for withdrawals not due to death	Matched Sham Set
	Adjustment of ranking criteria for JRT for permanent ventilation	Matched Sham Set
	Subjects who died will be excluded from the MI model	Matched Sham Set
	Rank based implementation of joint rank method + MI	Matched Sham Set

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

Table 7: Change from baseline to Day 183 in Motor Milestones total score: comparison: 50/28mg versus CS3B sham control

Type	Analysis Method	Analysis Population: Further Notes
	Joint rank using Jump to Control MI	Matched Sham Set
	Joint rank test using multiple imputation (MI), ANCOVA model not assuming equality of variance	Matched Sham Set

Endpoint and Comparison	Analysis Method	Analysis Population: further notes
Supplemental analysis		
Difference (50/28mg-Matched Sham Control) in change in HINE-2 total milestones from baseline by visit	ANCOVA using MI	Matched Sham Set
Difference (50/28mg-Matched Sham Control) in change in HINE-2 total milestones from baseline by visit	ANCOVA using MI (Not assuming equality of variances)	Matched Sham Set

6.4.3. Change in Plasma NF-L from (ratio to) baseline to Day 183: comparison: 50/28mg versus CS3B sham control.

Plasma Neurofilament

The analysis will be performed on the ITT set but is it noted that neurofilament plasma samples from China sites were not analyzed due to export law complication. The baseline and post baseline results for NF-L for these participants will be imputed in the MI step.

For Plasma NF-L, if the baseline is missing than the first available assessment post dose, within Day 2, will be used instead as baseline.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

Values that are BLQ will be set to half of the diluted LLOQ (2.6 pg/mL).

The Neurofilament light chain (NF-L) plasma concentration change (including absolute and ratio to baseline) will be presented by visit.

Summary tables and LS-mean (+- 95% CI) plots for geometric mean ratio to baseline using MI data +ANCOVA will be presented.

Main Analysis

- The change from baseline in Plasma NF-L from (ratio to) baseline to Day 183 will be analyzed using the JRT methodology to account for mortality. The ANCOVA model will include covariates for baseline log plasma NF-L, baseline CHOP INTEND total score and disease duration at screening. [redacted] Ranking based on death will be performed as described earlier for CHOP INTEND. In terms of ranking based on the NF-L, progressively higher ranks will be given to subjects with a greater reduction in NF-L at Day 183.

Sensitivity Analyses

- In order to test sensitivity of data handling for withdrawals not due to death, ranking criteria will be adjusted in the JRT for withdrawals not due to death.
- In order to test if permanent ventilation should be considered in evaluation of endpoint, ranking criteria will be adjusted for JRT to consider permanent ventilation in evaluation.
- The MI model will be run for total score for the ITT population excluding data from any subjects who died.
- Rank based implementation of joint rank method + MI

Supplementary analysis

- ANCOVA

For more details refer to section 6.3.

Table 8: Change in Plasma NF-L from (ratio to) baseline to Day 183: comparison: 50/28mg versus matched CS3B sham control

	Analysis Method	Analysis Population: Further Notes
Main analysis	Joint rank using multiple imputation (MI)	Matched Sham Set

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

Sensitivity analysis	Adjustment of ranking criteria for JRT for withdrawals not due to death	Matched Sham Set
	Adjustment of ranking criteria for JRT for permanent ventilation	Matched Sham Set
	Subjects who died will be excluded from the MI model	Matched Sham Set
	Rank based implementation of joint rank method + MI	Matched Sham Set

Endpoint and Comparison	**Analysis Method**	**Analysis Population: further notes**
Supplementary analysis		
Difference (50/28mg-Matched Sham Control) in change in Plasma NF-L from (ratio to) baseline by visit	ANCOVA using MI	Matched Sham Set

6.4.4. Change from baseline to Day 302 in CHOP INTEND total score: comparison: 50/28mg versus 12 mg Group from 232SM203.

Main Analysis

- The change from baseline to Day 302 in CHOP INTEND total score will be analyzed using the JRT methodology to account for mortality.

Sensitivity Analyses

- In order to test sensitivity of data handling for withdrawals not due to death, ranking criteria will be adjusted in the JRT for withdrawals not due to death.
- In order to test if permanent ventilation should be considered in evaluation of endpoint, ranking criteria will be adjusted for JRT to consider permanent ventilation in evaluation.
- The MI model will be run for total score for the ITT population excluding data from any subjects who died.
- Rank based implementation of joint rank method + MI.
- The main analysis will be repeated incorporating Jump to reference in the MI step.
- The main analysis will be repeated using PPS set.

Supplementary analyses

- ANCOVA

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

for more details section 6.3.

Table 9: Change from baseline to Day 302 in CHOP INTEND total score comparison: 50/28mg versus 12 mg Group from 232SM203

Type	Analysis Method	Analysis Population: Further Notes
Main analysis	Joint rank using multiple imputation (MI)	ITT Set
Sensitivity analysis	Adjustment of ranking criteria for JRT for withdrawals not due to death	ITT Set
	Adjustment of ranking criteria for JRT for permanent ventilation	ITT Set
	Subjects who died will be excluded from the MI model	ITT Set
	Rank based implementation of joint rank method + MI	ITT Set
	Joint rank using multiple imputation (MI)	PPS Set
	Main analysis repeated using Jump to reference in the MI step	ITT Set
	Joint rank using multiple imputation (MI)	ITT Set
	by disease duration	
	Log-rank test stratified by disease duration at Screening; by baseline CHOP INTEND	ITT Set

Endpoint and Comparison	Analysis Method	Analysis Population: further notes
Supplemental analysis		
Difference (50/28mg- 12 mg Group from 232SM203) in change in	ANCOVA using MI	ITT Set

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

CHOP INTEND from baseline by visit

6.4.5. Change from baseline to Day 302 in HINE Section 2 motor milestones total score: comparison: 50/28mg versus 12 mg Group from 232SM203.

Main, sensitivity and supplementary analyses

Main, sensitivity and s*upplementary* analyses will be analyzed in the same way as similar secondary endpoint described in 6.4.4.

Table 10: Change from baseline to Day 302 in HINE Section 2 motor milestones total score: comparison: 50/28mg versus 12 mg Group from 232SM203

Type	Analysis Method	Analysis Population: Further Notes
Main analysis	Joint rank using multiple imputation (MI)	ITT Set
Sensitivity analysis	Adjustment of ranking criteria for JRT for withdrawals not due to death	ITT Set
	Adjustment of ranking criteria for JRT for permanent ventilation	ITT Set
	Rank based implementation of joint rank method + MI	ITT Set
	Main analysis repeated using Jump to reference in the MI step	ITT Set

Endpoint and Comparison	Analysis Method	Analysis Population: further notes
Supplemental analysis		
Difference (50/28mg- 12 mg Group from 232SM203) in change in	ANCOVA using MI	ITT Set

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

HINE section 2 from baseline by visit	

6.4.6. Change in Plasma NF-L from (ratio to) baseline to Day 64: comparison: 50/28mg versus 12 mg Group from 232SM203.

Main and sensitivity analyses

The analysis will be performed on the ITT set but is it noted that neurofilament plasma samples from China sites were not analyzed due to export law complication. The baseline and post baseline results for NF-L for these participants will be imputed in the MI step.

Main and sensitivity will be analyzed in the same way as described in section 6.4.3.

Supplementary analyses

- Population level summary – ANCOVA

Table 11: Change in Plasma NF-L from (ratio to) baseline to Day 64: comparison: 50/28mg versus 12 mg Group from 232SM203.

Type	Analysis Method	Analysis Population: Further Notes
Main analysis	Joint rank using multiple imputation (MI)	ITT Set
Sensitivity analysis	Adjustment of ranking criteria for JRT for withdrawals not due to death	ITT Set
	Rank based implementation of joint rank method + MI	ITT Set

Endpoint and Comparison	Analysis Method	Analysis Population: further notes
Supplemental analysis		

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

Difference (50/28mg- 12 mg Group from 232SM203) in change in Plasma NF-L from baseline by visit	ANCOVA using MI	ITT Set

6.4.7. Time to death or permanent ventilation (≥ 16 hours ventilation/day continuously for > 21 days in the absence of an acute reversible event, or tracheostomy): comparison: 50/28mg versus CS3B sham control

Permanent ventilation is defined as tracheostomy or >=16 hours of ventilation/day continuously for > 21 days in the absence of an acute reversible event (Appendix B). Note that it was originally planned to include additional follow-up data from extension study in the analysis, but in order to focus this analysis plan on 232SM203 this will now be analyzed as part of the integrated 232SM203/232SM302 analysis plan.

An independent endpoint adjudication committee (EAC) will determine, in a blinded fashion, the date at which a participant is considered to have met the definition of an event. The procedures for reviewing and adjudicating event will be described in a charter.

The main analysis is to compare the time to death or permanent ventilation between the two treatment groups with the log-rank test stratified by disease duration at Screening (<=12 weeks or >12 weeks). Only events that were adjudicated by the EAC as meeting the protocol defined criteria for permanent ventilation or death will be included in the main analysis. Participants who do not meet the endpoint definition will be censored on the last occasion the participant was seen (either in-person visit or by telephone contact), irrespective of whether the participant has completed the full course of treatment and whether the participant has completed the study or permanently withdrawn. The exception is time to death or permanent ventilation cases in which a participant has begun a ventilator diary, in which case the latest entry in the diary will be used as the date of censoring.

Kaplan-Meier estimates of the cumulative probability of the time to death or permanent ventilation over time will be determined for the ITT population. Treatment comparisons of time to death or permanent ventilation will be based on a stratified log rank test. Stratification factor will be disease duration at Screening (<=12 weeks or >12 weeks). Kaplan-Meier plots will be presented. The median time to death or permanent ventilation, percentiles (5^{th}, 10^{th}, 25^{th}, 75^{th},

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

90th) and associated 95% confidence limits, and proportion of subjects who meet an event by fixed time points (i.e. Day 183 and Day 302) will be estimated using Kaplan-Meier.

The start date for calculating time to death or permanent ventilation in days i.e. death, permanent ventilation or censoring will be date of first dose. If a subject is adjudicated as having an event, the date they were considered to meet the endpoint either death or reach the definition of permanent ventilation will be provided as part of the adjudication.

The proportion of subjects meeting the endpoint of time to death or permanent ventilation will be presented by treatment group. For subjects who met the endpoint the number of days per subject at which ventilation use was at least 16 hours per day will be summarized.

Main analysis

Log-rank test stratified by disease duration at Screening.

Sensitivity Analyses

- The main analysis will be repeated using the actual treatment assignment. If participant, randomized to 12 mg nusinersen, receives 50 or 28 mg nusinersen at any point during the study, then the participant will be counted in the 50 mg nusinersen group.
- The treatment groups will be compared using a Cox proportional hazards model for time to death or permanent ventilation, adjusted for each subject's disease duration at Screening (rather than the binary stratification factor, ≤ 12 weeks/ > 12 weeks) and baseline CHOP INTEND total score. Only events that were adjudicated by the EAC will be included in this sensitivity analysis. This analysis allows estimation of a hazard ratio of nusinersen relative to the CS3B sham with its associated 95% confidence intervals.
- The main analysis will be repeated for the following subgroups: disease duration and Baseline CHOP INTEND

Table 12: Time to death or permanent ventilation: comparison: 50/28mg versus CS3B sham control

Type	Analysis Method	Analysis Population: further notes
Main Analysis	Log-rank test stratified by disease duration at Screening	Matched Sham Set

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

Sensitivity Analysis	Cox proportional hazards with treatment and each subject's disease duration at Screening and baseline CHOP INTEND total score as terms in the model	Matched Sham Set
	Log-rank test stratified by disease duration at Screening; by disease duration at Screening	Matched Sham Set
	Log-rank test stratified by disease duration at Screening; by baseline CHOP INTEND	Matched Sham Set

6.4.8. Time to death (overall survival): comparison: 50/28mg versus CS3B sham control

Overall survival will be analyzed in the same way as time to death or permanent ventilation.

Table 13: Time to death: comparison: 50/28mg versus CS3B sham control

Type	Analysis Method	Analysis Population: further notes
Main Analysis	Log-rank test stratified by disease duration at Screening	Matched Sham Set
Sensitivity Analysis	Cox proportional hazards with treatment and each subject's disease duration at Screening and baseline CHOP INTEND total score as terms in the model	Matched Sham Set
	Log-rank test stratified by disease duration at Screening; by disease duration at Screening	Matched Sham Set
	Log-rank test stratified by disease duration at Screening; by baseline CHOP INTEND total score.	Matched Sham Set

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

Table 13: Time to death: comparison: 50/28mg versus CS3B sham control

Type	Analysis Method	Analysis Population: further notes

6.4.9. Time to death or permanent ventilation (tracheostomy or ≥ 16 hours of ventilation/day continuously for > 21 days in the absence of an acute reversible event): comparison: 50/28mg versus 12 mg Group from 232SM203

Main and sensitivity analyses will be analyzed in the same way as similar secondary endpoint described in section 6.4.7.

Table 14: Time to death or permanent ventilation: comparison: 50/28mg versus 12 mg Group from 232SM203

Type	Analysis Method	Analysis Population: further notes
Main Analysis	Log-rank test stratified by disease duration at Screening	ITT Set
Sensitivity Analysis	Cox proportional hazards with treatment and each subject's disease duration at Screening and baseline CHOP INTEND total score as terms in the model	ITT Set
	Log-rank test stratified by disease duration at Screening; by disease duration	ITT Set
	Log-rank test stratified by disease duration at Screening; by Baseline CHOP INTEND	ITT Set
	Joint rank using multiple imputation (MI)	PPS Set

6.4.10. Time to death (overall survival) - comparison: 50/28mg versus 12 mg Group from 232SM203

Main, sensitivity analyses will be analyzed in the same way as similar secondary endpoint described in section 6.4.8.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

Table 15: Time to death: comparison: 50/28mg versus 12 mg Group from 232SM203

Type	Analysis Method	Analysis Population: further notes
Main Analysis	Log-rank test stratified by disease duration at Screening	ITT Set
Sensitivity Analysis	Cox proportional hazards with treatment and each subject's disease duration at Screening and baseline CHOP INTEND total score as terms in the model	ITT Set
	Log-rank test stratified by disease duration at Screening; by disease duration	ITT Set
	Log-rank test stratified by disease duration at Screening; by baseline CHOP INTEND	ITT Set
	Joint rank using multiple imputation (MI)	PPS Set

6.5. Supplemental analyses supporting primary and secondary endpoints.

6.5.1. Comparison: 50/28 mg Group versus CS3B sham control

The main analysis for the primary endpoint of CHOP INTEND: Change from baseline to Day 183 will be repeated for the Day 64 timepoint see table below for the specific analyses.

The Proportion of CHOP INTEND Responders at Day 183 will also be analyzed.

A subject is defined as a CHOP INTEND responder if the change from baseline in CHOP INTEND total score is greater than or equal to 4 points based on assessment at Day 183.

Subjects who die or withdraw from the study will be counted as non-responders and will be included in the denominator for the calculation of the proportion. Proportions of CHOP INTEND responders will be analyzed using logistic regression as described in the MI description in Section 6.1 and if due to low counts this is not possible then Fishers Exact test will be used.

Table 16: Supplemental Endpoints - comparison: 50/28mg versus CS3B sham control

Endpoint and Comparison	Analysis Method	Analysis Population: further notes

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

CHOP INTEND: Change from baseline to Day 64, accounting for mortality/dropout	Joint rank using multiple imputation (MI)	Matched Sham Set
Proportion of Motor Milestones responders at Day 64	Logistic regression or Fisher's Exact Test	Matched Sham Set
The Proportion of CHOP INTEND Responders at Day 183	Logistic regression or Fisher's Exact Test	Matched Sham Set

6.5.2. Comparison: 50/28mg versus 12 mg Group from 232SM203

For the comparison of higher dose and 12 mg Group from 232SM203, some of the secondary endpoints will be repeated as listed below, also the proportion of CHOP INTEND responders at Day 183 and Day 302 and proportion of Motor Milestones responders at Day 302 will be analyzed.

Table 17: Supplemental endpoints 50/28mg versus 12 mg Group from 232SM203

Endpoint	**Analysis Method**	**Analysis Population: further notes**
Change in CHOP INTEND from baseline to Day 183	Joint rank using multiple imputation (MI)	ITT set Repeat for Day 29 and 64.
Change in Motor Milestones total score from baseline to Day 183	Joint rank using multiple imputation (MI)	ITT set
The Proportion of CHOP INTEND Responders at Day 183	Logistic regression or Fisher's Exact Test	ITT set

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

The Proportion of CHOP INTEND Responders at Day 302	Logistic regression or Fisher's Exact Test	ITT set
Proportion of Motor Milestones responders at Day 302	Logistic regression or Fisher's Exact Test	ITT set

[redacted]

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

6.6. Analysis of the Remaining Secondary Endpoints - comparison: 50/28mg versus 12 mg Group from 232SM203

6.6.1. Change in CSF NF-L from (ratio to) baseline to Day 279

The analysis will be performed on the ITT set but is it noted that neurofilament CSF samples from China sites were not analyzed due to export law complication. The baseline and post baseline results for NF-L for these participants will be imputed in the MI step. Only baseline, Day 29 and Day 279 visits will be included as visits in the MI.

Values that are BLQ will be set to half of the diluted LLOQ (65 pg/mL).

The Neurofilament light chain (NF-L) CSF concentration change (including absolute and ratio to baseline) will be presented by visit.

Summary tables and LS-mean (+- 95% CI) plots for geometric mean ratio to baseline using MI data and ANCOVA will be presented.

6.6.2. Number and Length of Hospitalizations

The number of hospitalizations during the study will be analyzed using the rate at which they occur. For descriptive purposes, the aggregate hospitalization rate will be calculated for each group by dividing the total number of hospitalizations that occurred in the group by the total number of subject-years on study.

Annualized hospitalization rate will be calculated for each subject as the number of hospitalizations that the subject experienced divided by the number of days on study and this ratio multiplied by 365.25. Annualized hospitalization rate will be analyzed by a negative binomial regression model, adjusted for disease duration at screening and baseline CHOP INTEND total score. The logarithmic transformation of the time on the study will be included in the model as the "offset" parameter. If the data are under dispersed, or if the negative binomial regression model does not converge, a Poisson regression model with the same covariate will be used instead of the negative binomial regression model. Dispersion will be evaluated from the Pearson chi-squared statistic.

For each subject, the total time spent in hospital during the study will be calculated and prorated according to number of days on study. Based on the prorated time in hospital, treatment groups will be compared to an analysis of covariance adjusting for each subject's disease duration at Screening and baseline CHOP INTEND total score.

Analyses will be performed on the ITT Set.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

6.6.3. Clinical Global Impression of Change (CGIC) [Physician, Caregiver] at Day 302

CGIC will be assessed at Days 29, 64, 183 and 302. The CGIC rating scale was developed as a brief standalone assessment of the clinician's view of the participant's global functioning after initiating a study medication [Guy 1976]. The CGIC provides an overall clinician-determined summary measure that takes into account all available information, including knowledge of the participant's history, psychosocial circumstances, symptoms, behavior, and the impact of the symptoms on the participant's ability to function. The CGIC is administered by an experienced clinician who is familiar with the disease under study and the likely progression of treatment. The clinician makes a judgment about the total picture of the participant at each visit: the severity of the illness, the participant's level of distress and other aspects of impairment, and the impact of the illness on functioning. The CGIC is rated without regard to the clinician's belief that any clinical changes are or are not due to medication and without consideration of the etiology of the symptoms.

The CGIC is a 7-point scale that requires the clinician to assess how much the participant's illness has changed relative to a baseline state at the beginning of the intervention and is rated as follows: 1, very much improved; 2, much improved; 3, minimally improved; 4, no change; 5, minimally worse; 6, much worse; or 7, very much worse.

CGIC should be assessed consistently by the same rater for each study participant. A separate CGIC assessment will be performed by the Investigator (Principal Investigator or Sub investigator) and caregiver.

Clinical Global Impression of Change (CGIC) for the physician and caregiver at Day 302 will be analyzed as an ordered categorical score and as a dichotomized response: "much improved" or "very much improved" versus other responses.

The proportion of subjects with each CGIC score and with the dichotomized response of either "much improved" or "very much improved" will be summarized by visit and treatment group.

The difference in CGIC dichotomized response rate between 50/28 mg Group and Control Group will be analyzed using a logistic regression model with terms for treatment group and disease duration at screening and baseline CHOP INTEND total score. Participants who die or withdraw from the study, prior Day 302, will be counted as 'non responders'.

The odds ratio, its 95% CI and the p-value for 50/28 mg Group compared to the Control Group will be presented. The proportion of CGIC responders at other visits will also be analyzed in a similar way.

Analyses will be performed on the ITT Set.

6.6.4. Number of Serious Respiratory Events

All adverse events that are coded into the System Organ Class (SOC) of respiratory, thoracic, and mediastinal disorders, as their primary SOC will be considered respiratory events. Only

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

those classified as serious will be analyzed. By utilizing the above approach, rare occurrence of serious respiratory events in the Infection and Infestation SOC and Investigation SOC (if any) may not be included. However, a detailed review of the AE listing will be conducted to identify such events.

The number of serious respiratory events during the study will be analyzed using the rate at which they occur. For descriptive purposes, the aggregate rate will be calculated for each group by dividing the total number of serious respiratory events that occurred in the group by the total number of subject-years on study.

Annualized serious respiratory event rate will be calculated for each subject as the number of serious respiratory events that the subject experienced divided by the number of days on study and this ratio multiplied by 365.5.

Although serious respiratory events occur as a stochastic process and can be assumed to be from a Poisson distribution, studies in other therapeutic areas have shown that data of this type are typically over-dispersed, i.e., the variation seen is greater than that predicted by the Poisson model. To account for over-dispersion in the annualized rate of serious respiratory events, the negative binomial distribution will be used. Thus, on the basis of annualized serious respiratory event rate, treatment groups will be compared using a negative binomial model adjusting for disease duration at Screening and baseline CHOP INTEND total score. An "offset" parameter, the logarithm of the time on study, will be included in the model. If the data are under dispersed, or if the negative binomial regression model does not converge, a Poisson regression model with the same covariate will be used instead of the negative binomial regression model. Dispersion will be evaluated from the Pearson chi-squared statistic.

Analyses will be performed on the ITT Set.

6.6.5. Proportion of Time on Ventilation

The ventilation use during the study will also be collected via the daily ventilation diary. Daily ventilation use data that are missing will be imputed using the greater of the days that flank the missing day(s).

For each subject, the total number of hours of ventilation support required during the study will be calculated and prorated according to number of days on study. Based on the prorated number of hours of ventilation support, treatment groups will be compared via an analysis of covariance adjusting for each subject's disease duration at Screening and baseline CHOP INTEND total score.

An additional analysis of the number of days in which each subject required ventilation support of $\geq$ 16 hours per day will be performed. The number of days that this level of ventilation was required will be analyzed by analysis of covariance adjusting for each subject's disease duration at Screening and baseline CHOP INTEND total score.

Analyses will be performed on the ITT Set.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

6.6.6. Ventilator Use

The participant's ventilator use and respiratory aids will be collected at every study visit

The ventilation use will be summarized by visit and listed.

6.6.7. Change in the Parent Assessment of Swallowing Ability Scale (PASA)

Parent Assessment of Swallowing Ability (PASA) Dysphagia will be assessed using the PASA questionnaire. Caregivers will be asked a series of questions regarding the mealtime behavior of the participant.

The PASA questionnaire was developed by a Biogen team in order to assess the signs and symptoms of dysphagia. This questionnaire consists of 33 items across 4 domains that cover general feeding, drinking liquids, eating solid foods, and assessment of swallowing concerns. The first 3 of these domains are generally assessed with 5 levels of response (Never, Rarely, Sometimes, Often, and Always), although 2 items are assessed with a simple "Yes"/"No" answer. In the final domain, the assessment of swallowing concerns has 4 levels of response: Strongly Agree, Agree, Disagree, and Strongly Disagree. In answering each item, the caregiver is directed to consider the previous 7 days.

The number, mean score, and percentage of subjects scoring each category in the PASA Section 1 – General Feeding will be presented by visit. In addition, a stacked bar chart will be used to summarize the proportion of subjects in each category over time for each item. Depending on the number of subjects who complete PASA Sections 1,2, 3, and 4, similar presentations may also be made or alternatively the results from these visits will be listed.

7. Safety Endpoints

7.1.1. Adverse events

The analysis of safety will be performed using the respective Safety Sets and presenting by 232SM203 treatment groups. The sham participants from safety sets from CS3B will be included for the analysis of adverse events.

All AEs will be analyzed based on the principle of treatment emergence where treatment emergence will be relative to the first dose of nusinersen. An AE will be regarded as treatment-emergent if it was present prior to receiving the first dose of nusinersen and subsequently worsened in severity or was not present prior to receiving the first dose of nusinersen but subsequently appear. Refer to Appendix A for more details about AE with missing start or stop dates.

A participant having the same AE preferred term more than once will be counted only once in the incidence for that event. Multiple occurrences of the same adverse event for the same subject will all be counted in the frequency for that adverse event.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

The occurrence of the AE with the greatest severity will be used in the calculation of incidence by severity. The AE relationship to treatment or LP/sham procedure is captured as related or not related. In the event the relationship is missing then this will be considered related.

Due to the long half-life of nusinersen, analyses of treatment-emergent adverse events will include all events reported during the study. Adverse Events will be coded using the Medical Dictionary for Regulatory Activities.

The following presentations will be shown:

an overall summary showing, for each treatment group, the number and percentage of subjects with an adverse event, a mild, moderate or severe event, an event related to study drug, an event related to lumbar puncture /sham procedure, a serious event, a serious event related to study drug, an event leading to study drug withdrawal, an event leading to study withdrawal, and a fatal event.

incidence and frequency by system organ class and preferred term

incidence and frequency by preferred term

incidence and frequency, by preferred term, in at least 5% of subjects in any treatment group

incidence of adverse events by maximum severity by preferred term

incidence and frequency of severe events by system organ class and preferred term

incidence and frequency of events related to study drug by system organ class and preferred term

incidence and frequency of events related to lumbar puncture /sham procedure (as assessed by the investigator) by system organ class and preferred term

incidence by system organ class and preferred term that occurred within 24 hours of Dosing or Sham Procedure

incidence by system organ class and preferred term that occurred within 72 hours of Dosing or Sham Procedure

incidence and frequency of serious adverse events by system organ class and preferred term

incidence and frequency of serious adverse events related to study drug by system organ class and preferred term

incidence of events that led to fatal outcome

incidence of events leading to study drug withdrawal by system organ class and preferred term

incidence of events leading to withdrawal from study by system organ class and preferred term

incidence and frequency of adverse events over time by system organ class and preferred term.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

incidence and frequency of adverse events antibody status by system organ class and preferred term.

A separate table will also be presented to show an overall summary of COVID-19 pandemic related AEs.

Listings of the following events will be produced.

AEs

SAEs

AE with fatal outcome

AEs led to study drug withdrawal

AEs led to withdrawal from study

AEs related to lumbar puncture/sham procedure

AEs related to study drug

Follow-up adjusted incidence rate.

Follow-up adjusted incidence rate will be summarized by system organ class and preferred term by study group. Follow-up adjusted incidence rate – defined as the number of subjects who experienced an event divided by the total of entire follow-up time among the subjects in the analysis population. The entire follow-up time for subject is defined as the time from the first dose until the end of study day. Each subject will be counted only once within each category. A similar table will also be based on total number of events divided by the total of entire follow-up time among the subjects in the analysis population. This may count a subject more than once within each category of they experienced an event more than once.

Note on capture of severity and seriousness in CS3B.

In CS3B the eCRF captured detail for each investigator term for decreases in severity and changes from serious to non-serious. In 232SM203, the decrease in severity and change from serious to non-serious is not captured. In all studies worsening in severity and moving from non-serious to serious is captured.

In presenting CS3B as part of 232SM203 reporting, consistent with how these studies were originally reported in the following two scenarios:

1) where the severity decreases

2) where a serious event becomes non-serious

Only the first record will be considered treatment emergent so the actual reporting approach will be consistent with 232SM203.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

7.1.2. Concomitant therapy and Procedure

A concomitant therapy is any drug or substance administered between screening and the final study visit/telephone call. Participants are instructed to contact their Investigators before taking any new medications, including nonprescription drugs and herbal preparations.

Throughout the study, Investigators or designated licensed physicians involved in the study may prescribe concomitant therapies (procedures) or treatments deemed necessary for AEs or to provide adequate supportive care.

Participants are prohibited from receiving other experimental or approved agents for the

treatment of SMA, including gene therapy, biological agent, or device, during the study. This includes marketed agents at experimental doses that are being tested for the treatment of SMA (oral albuterol/salbutamol, riluzole, carnitine, sodium phenylbutyrate, valproate, hydroxyurea, etc.).

All concomitant therapies will be coded using the World Health Organization drug

dictionary (WHO Drug). A concomitant procedure is any therapeutic intervention (e.g., surgery/biopsy, physical therapy) or diagnostic assessment (e.g., blood gas measurement, bacterial cultures) performed between Screening and the final study visit/telephone call. Concomitant procedures will be coded using MedDRA.

For the purposes of analysis, a concomitant therapy (including medication or procedure) is defined as any therapy that was taken or administered on or after the first injection of nusinersen (see appendix A). This includes therapies and procedures that were started prior to the initiation of injection of nusinersen if their use continued on or after the first injection of nusinersen. The number and percentage of subjects who were taking each type of concomitant therapy and procedures at baseline and during the study will be presented by preferred term.

The number and percentage of subjects who underwent each type of concomitant procedure during the study will be presented.

Concomitant therapies related to LP/Sham procedure will be summarized separately.

If there are a sufficient number of tests or treatments reported with indication of COVID-19, a separate summary of these will be provided.

Concomitant medications will also be reviewed by a medical reviewer prior to the lock to determine disallowed medications according to the protocol. These will be summarized separately.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

7.1.3. Clinical laboratory data

Laboratory data will be evaluated to determine the incidence of abnormalities that emerge during the course of the study. Changes in laboratory evaluations will be presented relative to baseline, which is defined as the closest visit prior to the first dose.

The following clinical laboratory parameters are to be assessed:

Hematology panel: complete blood count with differential and platelet count (hematocrit, hemoglobin, platelets, red blood cell count [RBC], white blood cell count [WBC], basophils, eosinophils, lymphocytes, monocytes, neutrophils)

Blood chemistry panel: albumin, total bilirubin, direct and indirect bilirubin, alkaline phosphatase, alanine aminotransferase (ALT) (SGPT), aspartate aminotransferase (AST) (SGOT), gamma-glutamyl transferase (GGT), sodium, potassium, calcium, chloride, phosphate, blood urea nitrogen (BUN), creatinine, cystatin, creatine phosphokinase, creatine kinase, bicarbonate (CO2), glucose, total protein.

Urinalysis: specific gravity, pH, protein, glucose, ketones, bilirubin, blood, red blood cells, white blood cells, epithelial cells, bacteria, casts and crystals

Urine total protein assessed by local laboratories CSF analysis: RBC, WBC, protein, glucose.

Coagulation: PT (prothrombin time), aPTT (activated partial thromboplastin time), and INR (international normalized ratio).

Each hematology, blood chemistry, coagulation and CSF laboratory parameter will be flagged as “low”, “normal” or “high” relative to the parameter’s normal range or as “unknown” if no result is available.

For each urinalysis laboratory parameter, the number and percentage of subjects experiencing postdose shifts to abnormal will be summarized.

For each hematology, blood chemistry, coagulation and CSF parameter, the number and percentage of subjects experiencing postdose shifts to ‘low’ or ‘high’ will be summarized. In each summary, the denominator for the percentage is the number of subjects at risk for the shift. The number at risk for the shift to low is the number of subjects whose baseline value was not low and who had at least one post baseline value. The number at risk for the shift to high is the number of subjects whose baseline value was not high and who had at least one-post baseline value. Subjects will be counted only once for each parameter and each shift regardless of how many postdosing assessments had that type of shift. All post baseline data will be used in the shift tables, regardless of whether a scheduled or unscheduled visit.

Also, for each parameter, the incidence of shift to low will be summarized using the minimum post baseline values. Shift to low includes subjects with a normal, high, or unknown baseline value and at least one post baseline value of the given test. Similarly, the incidence of shift to

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

high will be summarized using the maximum post baseline values. Shift to high includes subjects with a low, normal, or unknown baseline value and at least one post baseline value.

To evaluate potential serious hepatotoxicity subjects with a post baseline AST and/or ALT value ≥3 times the upper limit of normal (ULN) and a post baseline bilirubin value >2 times ULN at any time, not necessarily concurrent, will be listed with their values. In addition, a plot will be presented with each subject's maximum post baseline AST or ALT value relative to the ULN against the subject's maximum post baseline bilirubin value relative to the ULN; values do not have to be concurrent. All post baseline data will be used to identify abnormalities, regardless of whether a scheduled or unscheduled visit.

Summary statistics for actual values and change from baseline in laboratory values will be summarized by treatment group and visit. Box plots for chemistry, hematology and CSF showing mean value for each treatment group at each visit will also be presented.

Due to the impact of the COVID-19 pandemic on transport infrastructure, public health or humanitarian emergencies, certain sites may only be able to test blood and urine samples at a local laboratory instead of central one. The use of local labs which will use different techniques/equipment and reference ranges may introduce variability in the results.

To assess the variability of the results two summaries will be presented for actual and change from baseline by visit:

- summarize local laboratory and central laboratory results together
- summarize only central laboratory results

Listings of all chemistry, hematology, coagulation, CSF, urinalysis values and pregnancy test (serum and urine) will be provided.

For platelets a summary of the proportion of participants with a post baseline platelet count below the lower limit of normal on at least 2 consecutive measurements will be presented.

For study 232SM203, the urine protein positive (trace, +1, +2, +3, +4 or >0.2g/L) results will be presented, number of subjects with positive result at baseline, with a negative result at baseline and at least one and at least two positive post baseline result respectively. This display will not be presented for CS3B and CS4 since the total protein from local laboratories was not collected per protocol.

7.1.4. Neurological Examinations

Hammersmith infant Neurological Examination (Sections 1 and 3) (HINE 1 and 3)

Sections 1 (neurological items) and 3 (behavior) of the HINE serve as the basis for neurological examinations. These are to be assessed at Screening, predose and at 3, 6, and 24 hours postdose on dosing days (Days 1, 15, 29, 64, 135, 183 and 279) and on Day 302/ET.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

The neurological items comprise cranial nerve function (facial appearance, eye appearance, auditory response, visual response, sucking/swallowing), posture (head in sitting position, trunk in sitting position, arms at rest, hands, legs in sitting position, legs in supine and standing positions, feet in supine and standing positions), movements (quantity, quality), tone (scarf sign, passive shoulder elevation, pronation/supination, adductors, popliteal angle, ankle dorsiflexion, pulled to sit, ventral suspension), and reflexes and reactions (tendon reflexes, arm protection, vertical suspension, lateral tilting, forward parachute). Behavior is comprised of state of consciousness, emotional state, and social orientation. For each item, for each subject, the worst post-baseline and the best post-baseline outcomes will be determined and 'shift' tables showing the shifts from baseline to the worst and from baseline to the best post-baseline value will be presented. In this analysis, all assessments post the first dose at baseline will be considering post baseline visits. For any item where the investigators have entered scores outside the set of expected values the scores will be mapped prior to determining shifts as follows: for baseline values, to the next highest expected value and for post-baseline values to the next lowest expected value.

Reflex assessments

Acute effects after dosing: Reflex

On each dosing day, each predosing reflex will be compared with all the post dosing assessments with 24 hours to see if the subject has changed (gained or lost) or maintained, reflexes. For this comparison the pre-dose reflex status will be categorized into two categories: 0 and >=1. For each reflex, visit, pre-dose category the number and proportion of subjects who change or maintained will be presented.

Acute effects after dosing: Reflex and other assessments

On each dosing day, the shifts from predose assessment on that day to postdose (i.e., from predose to 3, 6 and 24 hours postdose) will be determined and shifts the lowest post dose and the highest post baseline for each dose will be presented.

Chronic effects: Reflex and other assessments

For each item and subject, the lowest predose post baseline and the highest predose post baseline outcomes will be determined and 'shift' tables showing the shifts from baseline (predose value on Day 1) to the lowest predose post baseline and from baseline to the highest predose post baseline value will be presented. In this analysis, all assessments after the first dose will be considered post baseline visits.

For all the assessments the summary statistics for actual values and change from baseline in each parameter will be presented by treatment group and visit. Line plots for each item mean value for each treatment group at each visit and timepoint will also be presented.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

7.1.5. ECG

ECGs are to be recorded at Days 1, 15, 29, 135, 183, 279 and 302/ET, predose and post dosing at various time points on dosing days and end predose at 302/ET. The ECGs are assessed at a central reading laboratory and the results provided as external vendor data. On the eCRF the investigator interpretation is collected as normal, abnormal and abnormal AE and the investigator determination will be used in analyses.

The ECG test includes heart rate, QT interval, QTcF interval and RR interval. Summary statistics for actual values and change from baseline in each ECG parameter will be presented by treatment group and visit.

ECGs will be analyzed using two approaches.

Qualitative analysis

The determination of whether an ECG is abnormal not AE, abnormal AE made by the Investigator will be used for this qualitative analysis.

The number and percentage of subjects with shifts from normal to each of the categorical values denoting an abnormal scan (abnormal not AE, abnormal AE) will be summarized by treatment group. All post baseline data will be used to identify abnormalities, regardless of whether a scheduled or unscheduled visit.

Corrected Fridericia QT interval QTcF abnormalities (outliers)

QTcF will be examined to determine the incidence of clinically relevant abnormalities. The number of subjects evaluated and the number of subjects with clinically relevant abnormalities will be presented.

The first criteria for clinically relevant post dosing abnormalities (secondary safety endpoint) that will be considered is:

Post baseline QTcF of > 500 msec and Maximum increase from baseline to post baseline QTcF >60 msec.

Other criteria are:

Maximum increase from baseline QTcF > 30 to 60 msec

Maximum increase from baseline QTcF > 60 msec

Maximum post baseline QTcF > 480 to 500 msec

Maximum post baseline QTcF > 500 msec

ECG summaries will be presented for all subjects and for subjects with post baseline. Listings will include both the central and eCRF information.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

7.1.6. Vital Signs

The analysis of vital signs will be approached in two ways.

Vital signs are to be measured at Screening, predose and at 1, 2, 4, 6, and 24 hours postdose on dosing days (Days 1, 15, 29, 64, 135, 183 and 279) and on Day 302/ET. At each of these times, temperature, heart rate, respiratory rate, systolic and diastolic blood pressure, and pulse oximetry will be measured. If multiple evaluations are on the same visit day and visit time, then the mean of multiple evaluations will be selected for inclusion in the analysis.

Acute Effects After Dosing

On each dosing day, the change in each vital sign from predose on that day to postdose (i.e., from predose to 1 hour postdose, to 2 hours postdose, etc.) will be calculated. Summaries of actual values and change from predose for each dosing day will be presented.

Chronic Effects

To examine for possible chronic effects, the change from baseline (the predose value on Day 1) to the predose value on later dosing days and to Day 302/ET will be determined. Summaries of actual values and change from baseline for subsequent dosing days (Days 15, 29, 64,135, 183, and 279) and for Day 302/ET will be presented.

Vital signs (temperature, pulse, systolic and diastolic blood pressure and pulse oximetry) will also be examined to determine the incidence of potentially clinically relevant abnormalities.

Shift analysis

The number of subjects evaluated and the number of subjects with potentially clinically relevant abnormalities will be presented. All post baseline data will be used to identify abnormalities, regardless of whether a scheduled or unscheduled visit. The criteria for potentially clinically relevant postdose abnormalities are shown in Table 19 below.

Table 20: Criteria to determine potentially clinically relevant abnormalities in vital signs.

Vital Sign	Criteria for Abnormalities
Temperature	<36°C >38°C
Pulse	<60 bpm >100 bpm

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

Systolic Blood Pressure	<90 mmHg >140 mmHg >160 mmHg
Diastolic Blood Pressure	<50 mmHg >90 mmHg >100 mmHg
Weight	7% or more decrease from baseline 7% or more increase from baseline
Respiratory Rate	<12 breaths/min >20 breaths/min

7.1.7. Growth parameters

Growth parameters include length for age, weight for age, weight for length, head circumference for age, chest circumference, head to chest circumference ratio, and arm circumference.

The growth parameters are to be assessed during screening, on dosing days (Days 1, 15, 29, 64, 135, 183 and 279) and on Day 302/ET.

The weight for age and length for age will be summarized using the WHO child growth standards (WHO Child Growth Standards, 2006) for subjects aged up to 5 years and similarly the WHO growth reference data (WHO, 2007) for older subjects. These standards comprise percentiles for weight and length for age by sex and are available on a website (www.who.int) in txt files. Subjects will be cross referenced with these files, given the age and sex of the subject to determine below which percentile they lie for each parameter.

Summary statistics for actual values and change from baseline in each growth parameter will be presented by treatment group and visit.

Baseline growth parameters will be summarized and also the frequency and percentage for weight and length for age will also be presented for each of in the following categories: ≤ 1st, ≤ 3rd, ≤5th, ≤15th, ≤25th, ≤50th and > 50th percentile.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

7.1.8. LP opening pressure

LP opening pressure will only be measured at Days 1, 15, and 279 to avoid potential unblinding. A lack of published data exists in terms of the variability of opening pressure between and within subjects. It is suspected that a simple change in position of a subject during the measurement or change of instrument used contributes to variability in the measurement.

7.1.9. X-ray Examination of Spine

N/A

[redacted]

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

[redacted]

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

[redacted]

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

[redacted]

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

[redacted]

10. Pharmacodynamic Endpoints

As detailed earlier NF-L is defined as a secondary endpoint.

[redacted] In addition, plasma and CSF PK concentrations will be analyzed with plasma and CSF biomarkers to establish PK/PD correlation. PK/PD and exposure-response may also be conducted on efficacy. The results of this study may be combined with other nusinersen studies to perform population PK/PD analyses, including exposure-response analysis.

Other Analyses

Not applicable.

Statistical Considerations for Interim Analysis

Not applicable

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

Changes from Protocol-Specified Analyses

The secondary endpoint for plasma NF-L comparing Higher dose nusinersen to 12mg nusinersen was changed from Day 29 in the protocol to be Day 64 in this Statistical Analysis Plan

Summary of Changes from the Previous Version 1.0 of the SAP

Added in the per protocol set definition: Participants who had a previous diagnosis that would confound the data/conflict with SMA.

Added text to Section 6.4.1 to clarify that if baseline is missing, the median used in the imputation will be rounded up to the nearest integer value.

Added clarification to Section 6.4.3 that Ranking based on death will be performed as described earlier for CHOP INTEND. In terms of ranking based on the NF-L, progressively higher ranks will be given to subjects with a greater reduction in NF-L at Day 183.

Removed the analysis of CSF NFL with the joint rank test at Day 64 in Section 6.6.1. Day 64 is not feasible to use since the 50/28mg received a sham procedure at Day 64.

Two sensitivity analyses added to Section 6.1 based on comments from the FDA:

- Use of an ANCOVA model where the variance is estimated in each treatment arm separately for the main analyses of the change in CHOP INTEND and HINE-2 total score to Day 183 comparing sham to Higher dose.
- Implementation of Jump To Reference (J2R) and the joint rank test for the main analyses of CHOP INTEND and HINE2 total score for sham compared to higher dose and 12mg compared to higher dose.

Summary of Changes from the Previous Version 2.0 of the SAP

Based on comment from FDA added a sensitivity analysis for the main analyses using joint rank test of the change in CHOP INTEND and HINE-2 total score to Day 183 comparing sham to higher dose where the variance is estimated in each treatment arm separately.

References

[redacted]

Arkblad E, Tulinius M, Kroksmark AK, et al. A population-based study of genotypic and phenotypic variability in children with spinal muscular atrophy. Acta Paediatr. 2009;98(5):865-72.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

[redacted]

Berry JD, Miller R, Moore DH, et al. The Combined Assessment of Function and Survival (CAFS): a new endpoint for ALS clinical trials. Amyotroph Lateral Scler Frontotemporal Degener. 2013;14(3):162-8. Epub 2013/01/17.

Bohdana Ratitch, Ilya Lipkovich, Michael O'Kelly. Combining Analysis Results from Multiply Imputed Categorical Data. PharmaSUG 2013 - Paper SP03.

Chan, I. S. F., and Zhang, Z. (1999). “Test-Based Exact Confidence Intervals for the Difference of Two Binomial Proportions.” Biometrics 55:1202–1209. Coovert DD, Le TT, McAndrew PE, et al. The survival motor neuron protein in spinal muscular atrophy. Hum Mol Genet. 1997;6(8):1205-14.

De Sanctis R, Coratti G, Pasternak A, et al. Developmental milestones in type I spinal muscular atrophy. Neuromuscul Disord. 2016;26(11):754-759. Epub 2016/10/05.

Feldkötter M, Schwarzer V, Wirth R, et al. Quantitative analyses of SMN1 and SMN2 based on real-time lightCycler PCR: fast and highly reliable carrier testing and prediction of severity of spinal muscular atrophy. Am J Hum Genet. 2002;70(2):358-68.

Finkel RS, Mercuri E, Meyer OH, et al. Diagnosis and management of spinal muscular atrophy: Part 2: Pulmonary and acute care; medications, supplements and immunizations; other organ systems; and ethics. Neuromuscul Disord. 2017 Epub 2017/11/23.

Finkel RS, Mercuri E, Meyer OH, et al. Diagnosis and management of spinal muscular atrophy: Part 2: Pulmonary and acute care; medications, supplements and immunizations; other organ systems; and ethics. Neuromuscul Disord. 2018;28(3):197-207. Epub 2017/11/23.

Finkelstein DM, Schoenfeld DA. Combining mortality and longitudinal measures in clinical trials. Stat Med 1999; 18 (11):1341-54.f

Glanzman AM, Mazzone E, Main M, et al. The Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders (CHOP INTEND): test development and reliability. Neuromuscul Disord. 2010;20(3):155-61.

Glanzman AM, McDermott MP, Montes J, et al. Validation of the Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders (CHOP INTEND). Pediatr Phys Ther. 2011;23(4):322-6.

Guy W. ECDEU Assessment Manual for Psychopharmacology. Rockville, MD: U.S. Department of Health, Education, and Welfare; 1976.

Haataja L, Mercuri E, Regev R, et al. Optimality score for the neurologic examination of the infant at 12 and 18 months of age. J Pediatr. 1999;135(2 Pt 1):153-61.

ICH E9 (R1) Addendum 2017

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

Iannaccone ST, Hynan LS, Morton A, et al. The PedsQL in pediatric patients with Spinal Muscular Atrophy: feasibility, reliability, and validity of the Pediatric Quality of Life Inventory Generic Core Scales and Neuromuscular Module. Neuromuscul Disord. 2009;19(12):805-12.

[redacted]

Jedrzejowska M, Milewski M, Zimowski J, et al. Incidence of spinal muscular atrophy in Poland--more frequent than predicted? Neuroepidemiology. 2010;34(3):152-7.

Lefebvre S, Burlet P, Liu Q, et al. Correlation between severity and SMN protein level in spinal muscular atrophy. Nat Genet. 1997;16(3):265-9.

Little RJA, Rubin DB. Statistical Analysis with Missing Data. 2 nd ed. New York: John Wiley & Sons, Inc; 2002.

Main M, Kairon H, Mercuri E, et al. The Hammersmith functional motor scale for children with spinal muscular atrophy: a scale to test ability and monitor progress in children with limited ambulation. Eur J Paediatr Neurol. 2003;7(4):155-9.

[redacted]

Matsumoto H, Clayton-Krasinski DA, Klinge SA, et al. Development and initial validation of the assessment of caregiver experience with neuromuscular disease. J Pediatr Orthop. 2011;31(3):284-92.

Mazzone ES, Mayhew A, Montes J, et al. Revised upper limb module for spinal muscular atrophy: Development of a new module. Muscle Nerve. 2016 Epub 2016/10/04.

[redacted]

Mercuri E, Finkel RS, Muntoni F, et al. Diagnosis and management of spinal muscular atrophy: Part 1: Recommendations for diagnosis, rehabilitation, orthopedic and nutritional care. Neuromuscul Disord. 2018;28(2):103-115. Epub 2017/11/23.

[redacted]

O'Hagen JM, Glanzman AM, McDermott MP, et al. An expanded version of the Hammersmith Functional Motor Scale for SMA II and III patients. Neuromuscul Disord. 2007;17(9-10):693-7.

Permutt T, Li F. Trimmed means for symptom trials with dropouts. Pharm Stat. 2017;16(1):20-28. Epub 2016/08/15.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

Prior TW, Swoboda KJ, Scott HD, et al. Homozygous SMN1 deletions in unaffected family members and modification of the phenotype by SMN2. Am J Med Genet A. 2004;130A(3):307-10.

RBI: https://www.lshtm.ac.uk/research/centres-projects-groups/missing-data#dia-missing-data

Rosenbaum 2010, Design of Observational Studies

[redacted]

Schafer JL. Analysis of Incomplete Multivariate Data. 1 ed. London: Chapman & Hall/CRC Press; 1997. 430 p.

Schafer JL. Multiple imputation: a primer. Stat Methods Med Res. 1999;8(1):3-15

[redacted]

Varni JW, Seid M, Rode CA. The PedsQL: measurement model for the pediatric quality of life inventory. Med Care. 1999;37(2):126-39.

WHO Multicentre Growth Reference Study Group. Assessment of differences in linear growth among populations in the WHO Multicentre Growth Reference Study. Acta Paediatr Suppl. 2006;450:56-65.

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

Wirth B, Brichta L, Schrank B, et al. Mildly affected patients with spinal muscular atrophy are partially protected by an increased SMN2 copy number. Hum Genet. 2006;119(4):422-8.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

APPENDICES

APPENDIX A.

i Adverse Events

In the situation where worsening in severity (including seriousness) occurs for an adverse event, study sites are instructed to enter an end date and start a new record for the adverse event. In the new record, the changed severity is to be recorded and the start date will be set to be the end date of the previous record. Data linking those records are collected in the data base. Consider three scenarios:

- The first AE record occurs prior to the first dose, the second AE record occurs on or after the first dose, and the severity increases: Only the second record will be counted as treatment emergent.
- Both records occur on or after the first dose: If the AE severity on the second record is worse than the severity on the first record, then count both records as treatment emergent.
- Both records occur on or after the first dose: If the AE one is severe and the other serious then count both records as treatment emergent.

Of note, when counting the total number of treatment-emergent events, events linked through change in severity will still be counted as separate events.

For events with missing start or stop dates, the following criteria will be used for the purpose of identifying treatment-emergent adverse events:

If both the start and stop date for a particular event are missing, then the event is considered to have occurred on or after the first dose of study treatment.

If the start date for a particular event is missing and the stop date/time falls after the first dose date/time, then the event is considered to have occurred on or after the first dose of study treatment.

If the start time is missing and the start date is same as the first dosing date, then the event is considered to have occurred on or after the first dose of study treatment.

If it cannot be determined whether or not an event has occurred on or after dosing due to a missing or partial date, then the event will be assumed to have occurred on or after the first dose for the purpose of identifying treatment-emergent adverse events.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

Specifically, let AESTDT denote the start date of an adverse event and TRTSTDT be the start date of study treatment. For the purpose of identifying the treatment emergent adverse events, the following algorithm will be used for the imputation of missing or partial date:

If AESTDT is completely missing or the year is missing, then impute AESTDT to TRTSTDT.

If, in AESTDT, year is present and month/day are missing and year is equal to the year portion of TRTSTDT, then impute the month/day portion of AESTDT to the month/day portion of TRTSTDT.

If, in AESTDT, year is present and month/day are missing and year is not equal to the year portion of TRTSTDT, then impute the month/day portion of AESTDT to January 01.

Consider the situation in AESTDT where year and month are present with only day missing. If the year and month are the same as those for TRTSTDT, then impute day in AESTDT with day in TRTSTDT. Otherwise, impute the day in AESTDT with the first day of the month.

It is important to emphasize that the imputed date will not be used for calculations such as onset and duration of an adverse event.

ii Concomitant Therapy and Procedure

In order to define concomitant therapies with missing start or stop dates, the following additional criteria will be used:

- If both the start and stop dates of a particular therapy were missing, that therapy is considered concomitant.
- If the start date of a therapy was missing and the stop date of that therapy fell on or after the date of dosing, that therapy is considered concomitant.
- If the start date of a therapy was prior to the date of dosing and the stop date of that therapy was missing and the therapy was listed as continuing, that therapy is considered concomitant.
- If the start date of a therapy was prior to the date of dosing and the stop date of that therapy was missing and the therapy was listed as *not* continuing, that therapy is considered concomitant.
- If the start/stop date of a therapy is partial then where it is not possible to rule out that it was not taken concomitantly it will be considered concomitant.

Denote the end date of medications as CMENDT and the study treatment start date as TRTSTDT. The medication is classified concomitant provided any of the following is NOT true:

- CMENDT is complete and CMENDT is less than TRTSTDT.
- Day of CMENDT missing and year/month of CMENDT is strictly before year-month of TRTSTDT.

Docusign Envelope ID: 7DCA7D24-7D1B-4516-ABE7-DC0913457365

- Month of CMENDT is missing and year of CMENDT is strictly before year of TRTSTDT.

APPENDIX B.

PERMANENT VENTILATION DEFINITION CRITERIA: ACUTE REVERSIBLE EVENT

Purpose: Acute, intercurrent events may result in transient utilization of increased respiratory support for participants with spinal muscular atrophy (SMA) in this study. Such events may not reflect irreversible SMA disease progression and may complicate the ability to detect the effects of nusinersen on SMA disease progression.

A secondary endpoint is defined as the time to death or permanent ventilation (tracheostomy or ≥ 16 hours of ventilation/day continuously for > 21 days **in the absence of an acute reversible event**).

The purpose of this appendix is to define an acute reversible event and to indicate how this prespecified definition will be used in this study.

Any one of the following that occur between **7 days before and 7 days after** the onset of threshold-level respiratory support (≥ 16 hours/day) meet the definition of an acute reversible event:

1. Fever ≥ 102°F/38.9°C (tympanic, rectal, axillary, skin, sublingual)
2. Infection

Blood, sputum, throat, or cerebrospinal fluid (CSF) culture positive × 2 for virus, bacteria, or fungus

Blood, throat, sputum, or CSF viral polymerase chain reaction positive

Blood, throat, sputum, or CSF for infectious antigen diagnostic (e.g., streptococcus+ or hepatitis B surface antigen positive)

Blood, throat, sputum, or CSF positive for direct microscopic visualization (e.g., bronchoalveolar lavage or Gram stain revealing the presence of bacteria by tissue biopsy

DocuSign

Certificate Of Completion

Envelope Id: 7DCA7D247D1B4516ABE7DC0913457365
Subject: Complete with Docusign: 232SM203 Part B Infantile Onset SAP v3.0.docx
Vendor Name:
Source Envelope:
Document Pages: 83
Certificate Pages: 2
AutoNav: Enabled
EnvelopeId Stamping: Enabled
Time Zone: (UTC) Dublin, Edinburgh, Lisbon, London

Signatures: 3
Initials: 0

Status: Completed

Envelope Originator:
[redacted]
225 Binney St
Cambridge, MA 02142
[redacted]
IP Address: [redacted]

Record Tracking

Status: Original
19 July 2024 | 15:02

Holder: [redacted]

Location: DocuSign

Signer Events	Signature	Timestamp
[redacted] Principal Biostatistician Biogen Security Level: Email, Account Authentication (None)	DocuSigned by: [redacted] Signature Adoption: Pre-selected Style Using IP Address: [redacted]	Sent: 19 July 2024 | 15:05 Viewed: 19 July 2024 | 15:05 Signed: 19 July 2024 | 15:05
Electronic Record and Signature Disclosure: Not Offered via DocuSign		
[redacted] Spinal Muscular Atrophy (SMA) Biogen Security Level: Email, Account Authentication (None)	DocuSigned by: [redacted] Signature Adoption: Pre-selected Style Using IP Address: [redacted]	Sent: 19 July 2024 | 15:05 Viewed: 19 July 2024 | 15:06 Signed: 19 July 2024 | 15:06
Electronic Record and Signature Disclosure: Not Offered via DocuSign		
[redacted] Biogen Security Level: Email, Account Authentication (None)	DocuSigned by: [redacted] Signature Adoption: Pre-selected Style Using IP Address: [redacted]	Sent: 19 July 2024 | 15:05 Viewed: 19 July 2024 | 15:09 Signed: 19 July 2024 | 15:10
Electronic Record and Signature Disclosure: Not Offered via DocuSign		

In Person Signer Events	Signature	Timestamp

Editor Delivery Events	Status	Timestamp

Agent Delivery Events	Status	Timestamp

Intermediary Delivery Events	Status	Timestamp

Certified Delivery Events	Status	Timestamp

Carbon Copy Events	Status	Timestamp

Witness Events	Signature	Timestamp

Notary Events	Signature	Timestamp

Envelope Summary Events	Status	Timestamps
Envelope Sent	Hashed/Encrypted	19 July 2024 | 15:05
Certified Delivered	Security Checked	19 July 2024 | 15:09
Signing Complete	Security Checked	19 July 2024 | 15:10
Completed	Security Checked	19 July 2024 | 15:10

Payment Events	Status	Timestamps

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

PART B LATER-ONSET SMA STATISTICAL ANALYSIS PLAN

Version No.: 1.0

Date: 10April2024

Author: [redacted]

Study Title:
Escalating Dose and Randomized, Controlled Study of Nusinersen (BIIB058) in Participants With Spinal Muscular Atrophy

Name of Study Treatment: Nusinersen

Protocol No.: 232SM203

Study Phase: 2/3

Confidential Information

The information contained in this Statistical Analysis Plan (SAP) is confidential and may not be reproduced or otherwise disseminated without the approval of Biogen, Inc.

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

APPROVAL

This document has been reviewed and approved by:		
[redacted]	DocuSigned by: [redacted]	11-Apr-2024
SMT Statistician		Date
[redacted]	DocuSigned by: [redacted]	11-Apr-2024
Program Statistician		Date
[redacted]	DocuSigned by: [redacted]	11-Apr-2024
SMT Medical Director	Signature	Date

VERSION HISTORY

SAP Version	Date	Primary Reasons for Amendment

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

Table of Contents

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

LIST OF ABBREVIATIONS

Abbreviation	Definition
[redacted]	[redacted]
[redacted]	[redacted]
ACEND	Assessment of Caregiver Experience with Neuromuscular Disease
[redacted]	[redacted]
AE	Adverse event
ALT	Alanine Aminotransferase (SGPT)
ANCOVA	analysis of covariance
aPTT	Activated Partial Thromboplastin Time
AST	Aspartate Aminotransferase (SGOT)
AUC	Area under the curve
BLQ	Below limit of quantification
BUN	Blood Urea Nitrogen
CGIC	Clinical Global Impression of Change

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

CGIC	Clinical Global Impression of Change
CHOP INTEND	Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders
[redacted]	[redacted]
CO2	Carbon dioxide
COVID-19	Coronavirus disease 2019
CRF	Case report form
CSF	Cerebrospinal fluid
[redacted]	[redacted]
CV	Coefficient of variance
EAC	Endpoint adjudication committee
ECG	Electrocardiogram
GGT	Gamma-Glutamyl Transferase
HFMSE	Hammersmith Functional Motor Scale – Expanded
HINE	Hammersmith Infant Neurological Examination
HRQOL	Health-related quality of life
ICF	Informed consent form
ICH	International Council for Harmonization
IDMC	Independent data monitoring committee

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

INR	International Normalized Ratio
IRT	Interactive response technology
ITT	Intent-to-Treat
JRT	Joint Rank Test
LLOQ	Lower Limit of Quantification
LP	Lumbar puncture
L3M6	3 loading doses maintenance dose every 6 months
MI	Multiple imputation
NfL	Neurofilament Light chain
PASA	Parent Assessment of Swallowing Ability
PD	Pharmacodynamic(s)
PedsQL	Pediatric Quality of Life Inventory™
[redacted]	[redacted]
PMS	Post-marketing surveillance
[redacted]	[redacted]
[redacted]	[redacted]
PPS	Per protocol Set
PT	Prothrombin Time
QoL	Quality-of-life
QTcF	Corrected QT interval using Fridericia's formula

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

RULM	Revised Upper Limb Module
SAE	Serious adverse event
SD	Standard deviation
SE	Standard error
SMA	Spinal muscular atrophy
SMN	Survival motor neuron
SMN1	Survival motor neuron 1
SMN2	Survival motor neuron 2
SOC	System Organ Class
[redacted]	[redacted]
ULN	Upper Limit of Normal
US	United States
WHO	World A12:B65Health Organization

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

1. Introduction

This Statistical Analysis Plan is based on Version 6 of the protocol, dated 05May2022. All references to the protocol refer to Version 6.

Nusinersen is an antisense oligonucleotide administered intrathecally via lumbar puncture (LP); it increases survival motor neuron (SMN) protein expression and significantly improves motor function in patients with spinal muscular atrophy (SMA). Nusinersen was approved for the treatment of SMA under the tradename Spinraza™ in the United States (US), European Union, and other countries. The population for this study includes participants with infantile-onset and later-onset SMA.

Efficacy and safety results across the nusinersen clinical development program have demonstrated an overall positive benefit-risk profile of nusinersen across a broad range of SMA phenotypes and patient populations. Nusinersen is approved in the US, Europe, and other countries and regions for the treatment of SMA in paediatric and adult patients at a recommended dosage of 12 mg administered in 3 loading doses at 14-day intervals, a fourth loading dose 30 days after the third dose, and maintenance doses every 4 months thereafter in most countries and regions. Pharmacokinetic (PK) and pharmacodynamic (PD) analyses indicate that nusinersen drug exposure higher than that achieved with 12 mg in patients with SMA may produce an even greater benefit in motor function. Additionally, PK modelling and simulations identified dosing regimens that achieve higher drug exposure more rapidly. Therefore, this study is being conducted to investigate the efficacy, safety, tolerability, and [redacted] of a 50/28-mg dose of nusinersen (50-mg loading dose/28-mg maintenance dose) and a dosing regimen targeted to achieve higher drug exposure more rapidly. This study will be conducted in participants with genetically confirmed SMA.

This is a 3-part study Part A, Part B (infantile-onset and later-onset), and Part C.

Part A was an open label safety evaluation in which later-onset SMA subjects received 3 loading doses of 28 mg of nusinersen and 2 maintenance doses of 28 mg. All 6 participants enrolled in Part A completed.

Part B was double-blind, active-controlled controlled study designed to evaluate the proposed higher dosing regimen and included infantile and later-onset SMA participants.

Part C was an open label safety evaluation in which infantile and later-onset SMA of the participants, who were on the currently approved dose of nusinersen (12-mg maintenance for at least 1 year after the initiation of treatment) received higher dosing regimen via the administration of a single bolus dose of 50 mg of nusinersen (which should be administered 4 months ± 14 days after their most recent maintenance dose of 12 mg), with maintenance dosing at 28 mg thereafter. All 40 participants enrolled in Part C completed.

2. Study Overview

Study Objectives and Endpoints applicable to Part B later-onset.

This analysis plan it is only for Part B Later-onset analyses. Separate SAPs will cover the planned analyses for other parts of DEVOTE and the Part B infantile onset population.

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

The Part B primary objective is not applicable to the later-onset population.

Secondary Objectives

To examine the clinical efficacy of nusinersen administered intrathecally at higher doses compared to the currently approved dose in participants with SMA.

Secondary Endpoints

- Change from baseline in Hammersmith Functional Motor Scale – Expanded (HFMSE) score.
- Change from baseline in Revised Upper Limb Module (RULM) score.
- Total number of new World Health Organization (WHO) motor milestones
- Change from baseline in neurofilament light chain (NfL) concentration in plasma.
- Change from baseline in neurofilament light chain (NfL) concentration in CSF.
- Change from baseline in Assessment of Caregiver Experience with Neuromuscular Disease (ACEND)
- Change from baseline in Pediatric Quality of Life Inventory™ (PedsQL)
- Number and duration of hospitalizations.
- Clinical Global Impression of Change (CGIC) [physician, caregiver] at Day 302
- Number of serious respiratory events
- Ventilator use
- Change in the Parent Assessment of Swallowing Ability (PASA) scale.

Secondary Safety Objectives

To examine the safety and tolerability of nusinersen administered intrathecally at higher doses to participants with SMA.

Secondary Safety Endpoints

- Incidence of AEs, including SAEs
- Shifts from baseline in clinical laboratory parameters, ECGs, and vital signs
- Change in growth parameters
- Shifts from baseline in coagulation parameters (aPTT, PT, and INR)
- Change in urine total protein
- Change from baseline in neurological examination outcomes
- The proportion of participants with a post baseline platelet count below the lower limit of normal on at least 2 consecutive measurements
- The proportion of participants with a post baseline QTcF of > 500 msec and an increase from baseline to any post baseline timepoint in QTcF of > 60 msec

[redacted]

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

[redacted]

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

[redacted]

- [redacted]
- [redacted]

2.1. Study Design

This 3-part study will evaluate the efficacy and safety of a higher dosing regimen of nusinersen in approximately 145 participants. The study will be conducted at approximately 65 sites globally. Following the completion of this study, all eligible participants may elect to enrol in a separate long-term extension study (232SM302), pending study approval by ethics committees and the appropriate regulatory authorities. In regard to participants rolling over from Part B, this will be done without unblinding the participant's treatment group.

Part B will consist of a pivotal, double-blind, active-controlled treatment period with either 2 or 4 loading doses of nusinersen (for the 50/28-mg Group and Control Group, respectively) administered intrathecally followed by maintenance doses approximately every 4 months thereafter. Approximately 75 participants with infantile-onset SMA and 24 participants with later-onset SMA will be randomized in a 1:2 ratio to receive either the currently approved dosing regimen or a 50/28-mg dosing regimen.

In the Control Group, up to a total of 8 participants will receive 4 loading doses of 12 mg of nusinersen administered intrathecally (Days 1, 15, 29, and 64), followed by maintenance doses of 12 mg on Days 183 and 279 (and a sham procedure on Day 135).

In the 50/28-mg Group, up to a total of 16 participants will receive 2 loading doses of 50 mg of nusinersen administered intrathecally (Days 1 and 15) and 2 sham procedures on Days 29 and 64, followed by maintenance doses of 28 mg on Days 135 and 279 (and a sham procedure on Day 183). In order to maintain blinding sham procedures will be used as shown in the table below.

Arm	Dose 1	Dose 2	Dose 3	Dose 4	Dose 5	Dose 6	Dose 7
50/28-mg Group	D1 (50 mg)	D15 (50 mg)	D29 (sham)	D64 (sham)	D135 (28 mg)	D183 (sham)	D279 (28 mg)
Control (12/12 mg) Group	D1 (12 mg)	D15 (12 mg)	D29 (12 mg)	D64 (12 mg)	D135 (sham)	D183 (12 mg)	D279 (12 mg)

Randomization in Part B will be performed using interactive response technology (IRT).

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

For Part B later-onset participants, the randomization will be stratified by age at informed consent as follows:

< 6 years and

≥ 6 years

Blinded safety data will be reviewed on an ongoing basis by the Sponsor and the Medical Monitor. Unblinded safety data will be reviewed on an ongoing basis by an Independent Data Monitoring Committee (IDMC).

2.2. Sample Size Considerations

The sample size for Later-onset SMA in Part B is not based on statistical considerations.

A total of 24 participants with later-onset SMA will be randomized to the Control Group and 50/28-mg Group in Part B in a ratio of 1:2; this will allow the exploration of the safety, tolerability, [redacted], and efficacy of the 50/28-mg dose of nusinersen in this population. If any participants with later-onset SMA enrolled in Part B do not complete the Day 183 dose and assessments (e.g., due to withdrawal) as a result of a public health or humanitarian emergency, up to 4 additional participants may be enrolled at the Sponsor's discretion.

2.3. General Trial Conduct Mitigation Strategy under COVID-19 Pandemic or humanitarian emergencies

In order to mitigate risk of missing dosing or assessments during the time of the COVID-19 pandemic, the following has been put into place:

- If a site is not closed due to COVID-19 or humanitarian emergencies and the site enables the participant to attend the visit in-clinic, then delayed visits within a reasonable timeframe are allowed.
- Site transfers are encouraged where possible, so that participants can be transferred to another site for assessment. Instructions are in place to enable sites to do this including transfer of the database and source documents for the patient from the transferring site to the receiving site. This will allow the participant to receive dosing and protocol defined assessments to be made. Dosing can only be performed in the clinic setting due to intrathecal administration. In particular, if screening cannot be performed in the clinic, participants will be screened at a screening site and subsequently transferred to another site to perform dosing and all assessments for the participant, since dosing can only be performed in the clinic setting due to intrathecal administration.
- A phone visit can be used to collect safety data (AEs, concomitant therapies/procedures)
- Transferring sites are encouraged to enter and clean all data before transferring the participant to the receiving site. The database for a participant will remain with the transferring site until data are entered and cleaned as much as possible, but the transferring site must have access to the participant in IRT and copies of all source documents. Blinding of staff will be maintained as appropriate. Any queries raised later on for data belonging to visits at the transferring site will be cleaned by the receiving site through the source documents, and interaction between monitors at the receiving and transferring sites. The transferring site will no longer have access to the

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

database for a participant they have transferred but the monitors will. The Subject ID will remain the same when a participant is transferred to another site, but the transfer site can be tracked via the Site ID.

2.4. External Data

Three external datasets will be utilized to provide additional comparisons to the higher dose arm. These will include the CS4 study, CS11 (CS4 extension study) and the Japan PMS study.

2.4.1. CS4 Study

CS4 was a phase 3, randomized, double-blind, sham-procedure controlled study to assess the clinical efficacy and safety of nusinersen administered intrathecally in patients with later-onset spinal muscular atrophy. The study had a planned follow-up from first dose of 15 months.

Key eligibility criteria were onset of symptoms of SMA at > 6 months of age, aged between 2 and 12 years and ability to sit but never have achieved the ability to walk independently. The primary endpoint was the change in HFMSE score from baseline to Month 15 and secondary endpoints included the proportion of participants achieving a 3 point improvement on the HFMSE, change in RULM and proportion of participants achieving new WHO motor milestones.

The first patient was dosed in November 2014 and the last patient visit was February 2017. The study was stopped early based on the results of an interim safety and efficacy analysis and the study results were published (Mercuri et al. 2018).

A difference between CS4 and later-onset part B of 232SM203 was that in CS4, patients received 3 loading doses of 12 mg at days 1, 29, and 85 and then a dose at Day 274 in contrast to 232SM203 where 4 loading doses at Days 1, 15, 29 and 64 and then a dose at Day 183 is administered consistent with the nusinersen label in the majority of geographies.

2.4.2. CS11 Study

This was an open-label extension study in participants with SMA who previously participated in investigational studies of nusinersen. The last participant completed the study on 21Aug2023. The primary purpose of this study was to gather additional information on the long-term safety, tolerability, and efficacy of repeated doses of nusinersen. Participants who were randomized to receive sham control in CS4 switched to receive 12 mg nusinersen at the start of CS11, so this dataset provides additional patients who were naïve to treatment.

2.4.3. Japan PMS Study

This is an ongoing, observational study to survey all the SMA patients treated with nusinersen after the launch in Japan for the purpose of understanding the safety and effectiveness in Japanese patients. Patients are followed up to death, discontinuation of the treatment or the end of the study period (8 years after the launch), whichever occurs first. The dosing regimen approved in Japan for later onset participants is the 3 loading doses and 6 monthly maintenance dosing. As part of the effectiveness assessment, HFMSE has been collected at Day 274 dosing for later-onset patients. The latest data cut was at 30th of May 2023.

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

2.4.4. Summary of external data

The study design for 232SM203 included inclusion/exclusion criteria and efficacy assessments which are consistent with CS4, CS11 and the PMS study. Although the same dose of 12 mg is administered a number of differences exist and these are summarized in the following table.

Summary table of external data

	232SM203 Higher Dose	232SM203 Standard dose	CS4 nusinersen	CS4 Sham	CS11	PMS
Loading regimen	2 x 50mg doses	4 x 12mg doses	3 x 12mg doses		3 x 12mg doses	3 x 12mg doses
Maintenance regimen	28mg every 4 months	12 mg every 4 months	12mg every 6 months		12mg every 6 months	12mg every 6 months
Label for regimen		12mg L4M4	12mg L3M6		12mg L3M6	12mg L3M6
Timepoints efficacy collected (a)						
Month 6	Day 183	Day 183 [4]	Day 169 [3]	Day 169	NA	NA
Month 9	Day 279	Day 279 [5]	Day 274 [3]	Day 274	Day 265 [3]	Day 274 [3]
HFMSE collected	Yes	Yes	Yes	Yes	Yes	Yes
RULM collected	Yes	Yes	Yes	Yes	Yes	No
WHO motor milestones collected	Yes	Yes	Yes	Yes	Yes	No(b)
Estimated number of subjects	16	8	84	42	38(c)	27 (d)

a) Number in square brackets denotes the number of 12mg doses received prior to timepoint. So although in 232SM203 a dose is administered at Day 183 this is not counted in the 4 doses shown at Month 6.
b) WHO as collected in the PMS study uses different questions and is not considered comparable.
c) Subjects with baseline age 2 to <=10 years with HFMSE >=10 and <=54 and able to sit and never walked alone per WHO motor milestones.
d) Estimated based on data received Jan 2024. SMA type 2 patients, baseline age 2 to <=10 years with HFMSE >=10 and <=54.

2.4.5. Summary of How External Data Will be Used in the Analysis

In addition to the main comparison between the two active arms in study 232SM203, external data will be utilized for the following supportive analyses:

Comparison	How matching performed	Time point for efficacy comparison	Analyses
Higher dose versus sham control	[redacted]	Month 9, Month 6	Safety and efficacy (HFMSE, RULM, WHO, plasma NfL)
Higher dose versus nusinersen 3LM6 – CS4 only	[redacted]	Month 9, Month 6	Efficacy (HFMSE, RULM, WHO, plasma NfL)
Higher dose versus nusinersen 3LM6 – CS4 and CS11	[redacted]	Month 9	Efficacy (HFMSE, RULM, WHO)
Higher dose versus nusinersen 3LM6 – CS4, CS11, PMS study	[redacted]	Month 9	Efficacy (HFMSE)

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

Higher dose versus nusinersen 3LM6 – CS4 and CS11	[redacted]	Month 9	Efficacy (HFMSE, RULM, WHO)
Higher dose versus nusinersen 3LM6 – CS4	[redacted]	Month 9, Month 6	Efficacy (HFMSE, RULM, WHO, plasma NfL)
Higher dose versus sham control	[redacted]	Month 9, Month 6	Efficacy (HFMSE, RULM, WHO, plasma NfL)

3. Definitions

3.1. Dates and Points of Reference

- Study Day 1: the date of the first dose of study treatment in 232SM203
- Study Day
  - For a date on or after Study Day 1

    Study Day = (Date of Interest) – (Study Day 1) + 1
  - For a date before Study Day 1

    Study Day = (Date of Interest) – (Study Day 1)

In order to distinguish nominal visit names from duration defined in days, visit names will be referred to as “Day 15”, “Day 29”, etc., and “15 days” or “29 days”, etc. will be used to define time intervals.

Overall time on study will be defined as the total number of days a participant is known to be followed on study calculated as follows:

Overall time on Study = (Last date on study) – (Date of first dose) +1

Last date on study is defined as the date of the latest visit or evaluation or telephone contact, or time of death from all available data for a given participant.

Disease duration is defined as time from age at symptom onset to age at informed consent.

3.2. Study Treatments

The main presentations of data will be as follows:

Baseline characteristics, disposition, exposure, time on study and adverse event displays

232SM203		
12/12 mg nusinersen	50/28 mg nusinersen	Total

For baseline summaries including external data, all patients available in the pool will be included in the first column and then the matched patients, in the 2nd column.

CS4	Matched CS4	232SM203

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

Sham	Sham	50/28 mg nusinersen

For efficacy endpoints, safety other than AES, [redacted]:

232SM203	
12/12 mg nusinersen	50/28 mg nusinersen

For summaries of adverse events and efficacy including external data:

Matched CS4	232SM203
sham	50/28 mg nusinersen

The total study duration for each participant will be approximately 323 to 420 days:

- Screening: 21 days
- Loading period: 64 days
- Maintenance period: 215 days
- Follow-up: 23 to 120 days

3.3. Key Derived Variables

Not applicable.

3.4. Stratification Factors and Subgroup Variables

Randomization was stratified as follows:

- Age at informed consent:
  - < 6 years and
  - ≥ 6 years.

The subgroup categories are the following:

- One group comprising North America, Taiwan and Europe (excluding Russia) (regions contributing participants to original 12mg studies) versus another group comprising all other countries.

3.5. Analysis Sets

3.5.1. Within 232SM203

The Safety Set

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

The Safety Set is defined as all participants who receive at least one dose of nusinersen; participants will be analyzed in the treatment group based on what treatment they actually received.

The Intent-to-Treat (ITT) Set

The Intent-to-Treat (ITT) Set is defined as all participants who are randomized and receive at least one dose of nusinersen; participants will be analyzed in the treatment group to which they are randomized. This will be the primary population for the analysis of efficacy endpoints.

[redacted]

Per Protocol Set

A Per Protocol Set will include the subset of the ITT Set who have no significant protocol deviations that would be expected to affect efficacy assessments.

Significant protocol deviations will be determined prior to database lock and will include:

- Participants who had onset of signs and symptoms consistent with SMA at less than or equal to 6 months of age,
- who were younger than 2 years or older than 10 years,
- who cannot sit independently at screening,
- whose HFMSE is <10 or >54 at screening,
- who had respiratory insufficiency at Screening (defined as medical necessity for invasive or non-invasive ventilation for >6 hours during a 24-hour period) gtube at screening or severe scoliosis (Cobb angle > 40).
- Participants who took experimental or approved agents for the treatment of SMA, including gene therapy, biological agent, or device, during the study. This includes marketed agents at experimental doses that are being tested for the treatment of SMA (oral albuterol/salbutamol, riluzole, carnitine, sodium phenylbutyrate, valproate, hydroxyurea, etc.).

If participants have missed planned doses/sham but still completed study.

3.5.2. External data

The ITT Set as defined in the final analysis of CS4 will be used for the sham and nusinersen 12mg groups.

For CS11, the CS4 patients sham participants who received from the start of CS11 will be reassessed to determine if they would have qualified for 232SM203 based on fulfilling all the following conditions:

- Aged 2 to <=10 years at first dose
- Baseline HFMSE score ≥ 10 and ≤ 54

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

- Able to achieve WHO motor milestone of sitting without support and never achieved WHO walking alone.

Only patients who satisfy these criteria will be considered from CS11 in the matching pool.

For the PMS study a participant will need to meet all the following conditions to be included in the pool of participants:

- SMA type 2.
- Age at onset of symptoms > 0.5 years or missing.
- Aged 2 to <= 10 years at first dose.
- Baseline HFMSE score ≥ 10 and ≤ 54.

The following labels will be used to reference external data.

Label	Description
PS matched: CS4 Sham	Formed from the CS4 ITT set of sham control treated patients using a PS matching to select at least 16 comparable to the higher dose group
PS matched: nusinersen - CS4+CS11	Formed from the CS4 and CS11 ITT set of nusinersen treated patients using a PS matching to select at least 16 comparable to higher dose group
PS matched: nusinersen - CS4 only	Formed from the CS4 ITT set of nusinersen treated patients using a PS matching to select at least 16 comparable to higher dose group
PS matched with caliper: nusinersen A (Clinical trials and PMS)	12mg nusinersen participants from CS4, CS11 and Japan PMS study selected to match 232SM203 higher dose using [redacted]
PS matched with caliper: nusinersen B (Clinical trials)	12mg nusinersen participants from CS4 and CS11 studies selected to match 232SM203 higher dose using [redacted]
PS matched with caliper: CS4 Sham	CS4 sham participants selected to match 232SM203 higher dose using [redacted]

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

4. List of Planned Study Analyses

4.1. Interim Analysis

No interim analysis is planned. Within Part B, when later-onset patients are completed then the final analysis will be performed on the later-onset patients. The unblinding plan details the steps to ensure that the blind is maintained throughout the entire study and for the infantile-onset patients should the later-onset patients be unblinded.

4.2. Primary Analysis

N/A

4.3. Final Analysis

The final analysis will be performed when all later-onset participants in Part B have completed the follow-up period.

4.4. General Principles

Descriptive summary statistics will be presented for all endpoints collected. Unless otherwise specified, for continuous endpoints, the summary statistics will generally include number of participants with data, mean, standard deviation, median, 25th percentile, 75th percentile, minimum and maximum. For categorical endpoints, the summary statistics will generally include number of participants with data, and the percentage of those with data in each category.

All summaries and listings will be presented by study groups unless otherwise specified. Visits in listings will be displayed as per CRF data collection rather than analysis visits.

The statistical software, SAS®(Version 9.4) will be used for all summaries and statistical analyses.

Visit Windows for Early Withdrawal, Unscheduled Visits

Data from early withdrawal visits and post baseline unscheduled assessments will be assigned to an appropriate scheduled post baseline visit using a windowing scheme for assessments that are tabulated or summarized by visit. Scheduled visits will not be windowed.

The visit windowing will be performed for efficacy and safety endpoints.

For safety endpoints the following windows will be utilized.

Table 2: Visit windows for safety Endpoints in 232SM203

Visit	Lower Bound	Upper Bound	Target Day
Baseline			<=1
Day 15	2	21	
Day 29	22	47	29
Day 64	48	99	64

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

Day 135	100	159	135
Day 183	160	232	183
Day 279	233	290	279
Day 302	291	348	302

For efficacy endpoints, the actual day of the unscheduled or end of study visits will be considered.

For post baseline visits up to and including Day 64 a +/- 7 day window will be used and for visits between Day 135 and Day 302 a +/- 14 day window will be used.

The following rules will be implemented for windowing:

- If more than one observation is within the same window, data from the regular scheduled visit will be used for that visit.
- If neither of the observations are from a regular scheduled visit the observation closest to the planned target date will be used. However, if both the observations are equidistant from the target date, the latest one will be used; if repeated measurements are on the same day, then the last measurement will be used.
- If windowing safety data safety data and there is more than one observation in a window for a dosing visit the observation on the day of dosing will be chosen over the observation where dosing did not occur on the same day.
- In windowing laboratory data, if a central and local result is available within the window then the central laboratory result will be chosen.

Summary table of windowing for the external data

Study	Windowing
CS4	All dosed subjects will be included, and the Month 6 and Month 9 values collected will be utilized.
CS11	All dosed subjects meeting the re-evaluated conditions will be included. The protocol defined Day 265 visit will be used. If this was not collected since the subject moved to MMDR dosing, then Day 265 will be missing.
PMS study	The HFMSE visit data will be windowed to give Month 9 as study day 274 +/- 15 days. Therefore, study days between 259 and 289 will be considered, if >1 lie in the window the one closest to Day 274 will be selected. If two are equally close, the latest one will be selected.

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

4.5. Participant Accountability

The number (and percentage) of participants screened, screening failure, randomized, withdrew prior to dosing, dosed, completed treatment, discontinued treatment and the reasons for discontinuation from treatment, completed the study, withdrew from the study and reason for withdrawal will be summarized in a table.

A listing of participants who discontinued treatment/withdrew from study and the associated reasons for discontinuation/withdrawal will be presented. A separate listing will also be presented for participants who discontinued treatment/withdrew from study due to the COVID-19 pandemic or to humanitarian emergencies if applicable.

4.6. Demographic and Baseline Characteristics

Baseline data (demography, medical history, SMA history, and baseline disease characteristics) will be summarized.

Formal statistical analyses will not be done to test for homogeneity between treatment groups. If there are apparent heterogeneities between the groups in any of the participant characteristics that are of clinical importance or could affect the treatment outcome, the impact of the imbalances will be investigated and, if appropriate, adjustments made in the efficacy and safety analyses.

Demography includes age at screening, age at first dose, sex, ethnicity, race, country and geographic region. Medical history will be coded in MedDRA and the number and percentage of participants with each medical history presented by preferred term.

SMA history includes age of symptom onset, time from disease onset to enrollment, age at SMA diagnosis, time from diagnosis to enrollment, number of copies of the SMN1 and SMN2 gene, highest motor function achieved and maintained, wheelchair use.

Baseline disease characteristics will be assessed by HFMSE, WHO motor milestones, Revised Upper Limb Module Test Total score and Entry item, growth parameters and Cobb angle. These assessments are further described in the next paragraph and Section 5.

Demographic and baseline disease characteristics will be presented for the ITT Set, the PPS and the Safety Set as appropriate.

4.7. Protocol Deviations

Protocol deviations identified during site monitoring will be captured in a protocol deviation log and categorized as major or minor deviations. These deviations will be listed. Major protocol deviations will be summarized including those that lead to exclusion from the Per protocol Set.

Participants with incorrect stratification i.e. any mismatch between the stratification used for randomization on the IRT system and the actual stratification according to the eCRF will also be listed. A separate summary of major protocol deviations related to the COVID-19 pandemic will be presented.

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

4.8. Study Treatment Exposure

The number of doses received, and the number of sham procedures performed will be displayed using frequency distributions. The cumulative amount of nusinersen received will be summarized using summary statistics.

The overall time on study will be summarized descriptively by treatment. Overall time on study will be calculated (section 3.1). The duration will also be categorized and summarized using the following categories: <29 days, >=29 to 64 days, >=64 to 99 days, >=99 to 135 days, >=135 to 183 days, >=183 to 218 days, >=218 to <260, >=260 to 302 and >=302 days

Given the long half-life of nusinersen, participants are considered to be exposed from the time the first dose of nusinersen was administered (in or before the start of study) to the date of last visit or contact.

Separate listings will be provided showing what participants were randomized to as well as study drug administration data which will include lot numbers, actual treatment received, cumulative number of doses and cumulative doses. A listing of dosing errors will also be provided i.e. where either the incorrect dose was administered compared with the kit description from IRT system. A listing will also be presented to show whether any of the missed doses are related to COVID-19. If any of the doses are due to COVID-19, then the number of impacted doses will be presented.

5. Efficacy Endpoints

5.1. General Analysis Methods for Efficacy Endpoints

5.1.1. Multiple imputation (MI)

In the event of missing data for post baseline visits then multiple imputation will be performed.

For secondary, [redacted] endpoints the multiple imputation method [Schafer 1997, Schafer 1999] will be used for patients who discontinue for a reason other than death or have a missing efficacy result at the timepoint of interest. This will be performed for HFMSE, RULM, NfL CSF and plasma concentration, ACEND, and PedsQL.

The imputations will be performed on total scores for post baseline visits. All available data will be used i.e., where participants have missed doses but continued to perform assessments; otherwise, any missing values will be imputed using MI.

If the MI procedure imputes any values outside the expected range for the scale e.g. from 0 to 66 for HFMS-E then values below 0 will be set to 0 and values above 66 will be set to be 66, and for NfL the minimum will be set to LLOQ/2.

The Markov Chain Monte Carlo (MCMC) method will be used to impute the missing score.

The treatment variable will be coded so that the Higher dose group is the first in the sort order. Prior to the Proc MI step the dataset will be sorted by treatment and then ascending unique subject identifier (USUBJID). The variable list in the model for imputations will include:

- treatment group, age at screening, baseline plasma NfL baseline value for the endpoint, and all available post baseline values

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

A set of 100 complete imputed datasets will be generated and the relative efficiency parameter will also be checked to determine the acceptability of the imputed results i.e. it should be close to 98% or higher.

For each of the 100 imputed datasets, the endpoint will be compared between treatment groups using an ANCOVA model for continuous endpoints or logistic regression for binary endpoints as described in the main analysis for each endpoint.

The estimates from the 100 fitted models will be combined to provide an overall estimate with corresponding confidence intervals and p-value using PROC MIANALYZE [Little et al, 2002].

For the continuous endpoints, the difference between treatments and the corresponding CI will be presented. For the analyses using logistic regression the odds ratio and corresponding CI will be presented. In addition, the estimates of the binomial proportions in each treatment arm and the differences between treatments will be presented (Ratitch, 2013).

In the presentation of results from multiple imputed data, the number of subjects with missing data will be summarized.

For any given endpoint such as HFMSE, a number of different comparisons are planned such as within study 232SM203: 12mg to higher dose, sham versus higher dose, L3Q6 compared to higher dose. For each comparison a separate multiple imputation procedure will be performed including the 2 corresponding treatment groups. Only visits in common between the two arms will be included in the multiple imputation step. This will mean that for any comparisons including CS4, CS11 and the PMS study, only baseline and Month 9 HFMSE visits will be included in the imputation step and for comparisons including CS4 only, baseline, Month 6 and Month 9 will be included.

5.1.2. Matching external data

Due to the wide range of matching algorithms available and the difficulty in foreseeing all challenges upfront an additional operational step will be implemented as follows as part of the comparisons to external data:

An independent statistician will be identified. This statistician will be considered independent as they will have no knowledge or ability to access Part B later-onset post-baseline data in DEVOTE.

- The independent statistician will be provided with:232SM203 baseline covariates, baseline efficacy data and be unblinded to he assigned treatments. They will not have access to any on-treatment efficacy data for 232SM203Baseline data from CS4, Japan PMS, CS11.
- The following process will be followed:

1) The independent statistician will be provided with code to perform matching per this analysis plan and attempt to find the best match they can per the matching designs. In a situation where the matching approach is not successful then the statistician may implement an alternative matching algorithm.
2) If a match is found, then the independent statistician will create a dataset which identifies which groups of subjects were matched between studies.

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

3) Following the unblinding of 232SM203 the statistician responsible for the efficacy analysis will be provided with the file in (2) and perform the planned efficacy analysis using the groups of subjects identified. Since this statistician will have access to post-baseline efficacy data, they will not be permitted to perform the matching step (1,2).

This operational sequence is considered important in order to increase the chance of making a reasonable match, but also to protect the integrity of matching and minimize the bias which could occur if the statistician performing the matching had access to the efficacy data (or vice versa). Further details on this are described in the unblinding plan.

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

[redacted]

Matching designs

1. Comparison of higher dose versus 12 mg nusinersen L3M6 dosing regimen

The intent is to match the 16 subjects in 232SM203 who received higher dose with a subset of subjects selected from the pool of subjects who received 12mg from baseline in: CS4, CS11 or the Japan PMS study. The matching design will follow the steps and number of subjects outlined in the following table:

Design	232SM203 higher dose	CS4 or CS11 or Japan PMS 12 mg

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

1) Match 1 to 3	N= 16	N= 48
2) Match 1 to 2	N= 16	N= 32
3) Match 1 to 1	N= 16	N=16

2. Higher dose versus 12 mg nusinersen from CS4 and CS11 controlled clinical trials L3M6 dosing regimen.

The matching design will be the same design as the design for comparison 2 but will use a smaller pool of patients as Japan PMS patients will not be included. In this analysis, baseline RULM will also be included as a covariate for matching [redacted].

3. Matching design for comparison 4: higher dose versus CS4 Sham control

The matching design will be as shown in the table below. In this analysis baseline RULM will also be included as a covariate for matching [redacted].

Design	232SM203 higher dose	CS4 Sham
1) Match 1 to 2	N= 16	N= 32
2) Match 1 to 1	N= 16	N= 16

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

5.1.4. Primary Efficacy Endpoints

N/A

5.1.5. Secondary Efficacy Endpoints

The following secondary efficacy endpoints will be evaluated for later-onset SMA population between 50/28 mg Group and 12/12 mg Group:

- Change from baseline in Hammersmith Functional Motor Scale – Expanded (HFMSE) score
- Change from baseline in Revised Upper Limb Module (RULM) score
- Total number of new World Health Organization (WHO) motor milestones
- Change from baseline in neurofilament light chain (NfL) concentration in plasma
- Change from baseline in neurofilament light chain (NfL) concentration in CSF
- Change from baseline in Assessment of Caregiver Experience with Neuromuscular Disease (ACEND)
- Change from baseline in Pediatric Quality of Life Inventory™ (PedsQL)
- Number and duration of hospitalizations
- Clinical Global Impression of Change (CGIC) [physician, caregiver] at Day 302
- Number of serious respiratory events
- Ventilator use
- Change in the Parent Assessment of Swallowing Ability (PASA) scale

Analysis Methods for HFMSE

The HFMSE is a tool used to assess motor function in children >2 years or older with later-onset SMA. The scale was originally developed with 20 scored activities and was devised for use in children with Type 2 and Type 3 SMA with limited ambulation to give objective information on motor ability and clinical progression. The expanded scale includes an additional module of 13 items developed to alleviate the ceiling effect and allow evaluation of higher functioning and ambulatory SMA patients. Each item is scored 0 (unable), 1 (performs with modification or adaption) or 2 (able) and the total score is calculated by summing the 33 items and ranges from 0 to 66 with higher scores indicating greater motor function. If 6 or fewer items are missing, then these items will be imputed to 0 when summing all 33 items. If greater than 6 items are missing, then the total score will be set to missing.

Baseline is defined as the last non-missing total score prior to the first dose of nusinersen.

In order to further explore the response as measured by HFMSE score a number of thresholds will be evaluated: The proportion of subjects achieving: worsening of >= 4, >= 3, >=2, >=1 points, any worsening, no change (=0 points), any improvement, >=1, >=2, >=3,>=4 points improvement will be presented.

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

The main analysis is to compare the change from baseline HFMSE score at Day 302 between 50/28 mg Group and 12/12 mg Group using an ANCOVA model with treatment group as a factor and age at screening, baseline Plasma NfL and baseline HFMSE score as covariates. Missing post baseline HFMSE data will be handled using the multiple imputation method.

Several sensitivities will be performed for change from baseline HFMSE:

- The main analysis will be repeated using the PPS.
- The analysis will be performed using a Mixed Effect Model with Repeated Measures (MMRM) with an unstructured covariance matrix, where the treatment group, time (as a categorical variable), treatment by time interaction, and age at screening, baseline Plasma NfL, baseline HFMSE and baseline by time interaction will be included as covariates. No imputation for missing visit data will be performed and the analysis will be performed on the ITT set.
- The missing Day 302 values will be imputed by "last observation carried forward". This will be performed on the ITT set.
- Multiple imputation will be repeated, but for subjects who have a missing Day 302 assessment due to discontinuation for any reason, the worst of the subjects last observed value or the baseline value will be imputed. For subjects who completed but have a missing Day 302 value the imputed values will be used.

In order to further explore the change in HFMSE score during the study and the difference between treatment groups the change from baseline to each visit will be analyzed as described for the main analysis. A plot of mean change from baseline (based on the estimates from the ANCOVA model) over time will be provided by treatment group.

Proportion of subjects achieving>=3 point improvement in HFMSE from baseline at Day 302:

The multiple imputation datasets produced for the main analysis will be utilized. If a subject terminates the study prior to Day 302 assessment, then the imputed value from the multiple imputation will be ignored and the subject will be considered a non-responder. Any subject in the ITT Set who has an increase from baseline HFMSE score of 3 or more points at Day 302 will be defined as a responder; otherwise, a subject will be considered a non-responder. Treatment groups will be compared using logistic regression adjusting for each subject's age at screening and baseline HFMSE.

If the logistic regression model does not converge then a Cochran–Mantel–Haenszel test will be performed instead.

Analysis Methods for RULM

The upper limb module test (Mazzone et al. 2011) was developed to assess the upper limb functional abilities in SMA patients using 9 items. However, a revised version of the upper limb module test developed by the Upper Limb Module Working Group will be performed in this

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

study consisting of 20 items. The first item is assessed on a seven point scale ranging from 0 – "No useful function" to 6 – "Can abduct both arms simultaneously elbows in extension in a full circle until they touch above the head". This first item will not contribute to the total score. For the remaining 19 items, 18 items are assessed on a three point scale using the following criteria: 0 – Unable to achieve independently; 1 – Modified method but achieves goal independent of physical assistance from another person; 2 – Normal – achieves goal without any assistance. The one remaining item is assessed as either 0 or 1. For each item, a score will be collected on the left and right side. A derived total score will be calculated by summing the scores from these 19 individual items and ranges from 0 if the subject fails all activities to 37 if the subject achieves all activities. If, for an individual item, a response is recorded for both he left and right side the highest score will be used in calculating the total. If 3 or fewer items are still missing responses, then it will be assumed that the score was 0. If greater than 3 items are missing, then the total score will be set to be missing.

Baseline is defined as the last non-missing total score prior to the first dose of nusinersen.

In order to further explore the response as measured by upper limb score a number of thresholds will be evaluated: The proportion of participants achieving worsening of >= 4, >= 3, >=2, >=1 points, any worsening, no change, any improvement, >=1, >=2, >=3,>=4 points improvement will be presented.

The main analysis is to compare the change from baseline RULM total score at Day 302 between 50/28 mg Group and 12/12 mg Group using an ANCOVA model with treatment group as a factor and age at screening, baseline Plasma NfL and baseline RULM score as covariates. Missing post baseline RULM data will be handled using the multiple imputation method. Sensitivity analyses as described for the HFMSE will be performed.

Analysis Methods for Total Number of New WHO Motor Milestones

The WHO motor milestones are a set of six milestones in motor development, all of which would be expected to be attained by age 24 months in healthy children. The individual milestones are:

- Sitting without support
- Standing with assistance
- Hands and knees crawling
- Walking with assistance
- Standing alone
- Walking alone

Per the study inclusion criterion number 16, subjects would be expected to all have achieved and maintained at baseline the milestone "Sitting without support" but none would have achieved the milestone of "Walking alone". The motor milestones are assessed using the WHO motor milestone criteria [WHO Multicentre Growth Reference Study Group 2006]. As part of the

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

assessment, the examiner records an overall rating of the subject's emotional state and then for each milestone one of the following four classifications:

- No (inability) – Child tried but failed to perform the milestone
- No (refusal) – Child refused to perform despite being calm and alert
- Yes – Child was able to perform the milestone
- Unable to test – Could not be tested because of irritability, drowsiness or sickness.

Imputation for motor milestones

If for a milestone either "No (refusal)" or "Unable to test" are observed at a visit, then the result will be first set to missing. Imputation will be performed for missing data considering each milestone separately using the following rules for scheduled visits.

For baseline, the closest non-missing milestone prior to or at first dose will be selected. If the motor milestone is still missing, then the missing value will be imputed as the median of the non-missing values of the stratum to which the subject belongs to: Age at screening <6 years or Age at screening >=6 years. In the event that the median at baseline is 0.5 then the missing value will be imputed as 1. If, for the subject with a missing value at a particular visit the corresponding visit is flanked by visits with non-missing milestones, the missing value will be imputed by using the worst result from the flanking visits. Otherwise, if the imputation is the last visit, the missing value will be imputed as the lowest value observed across all subjects who have a non-missing value at this visit within the stratum: Age at screening <6 years or Age at screening >= 6 years within the treatment group. Of note, only observed data will be utilized for imputation purposes and in situation where unscheduled visit data is available these values will be utilized if these flank a missing visit. Missing motor milestone items will be imputed first prior to any analysis.

Main Analysis

The main analysis is to compare the total number of new motor milestones achieved at Day 302 between 50/28 mg Group and 12/12 mg Group, based on the ITT set. Missing data will be handled as described above and a total score for each visit calculated by summing the six milestones.

For each milestone the Day 302 and baseline values will be compared and if the subject can achieve a new milestone compared to baseline, then it will be counted +1, if the subject is unable to achieve the milestone, then it will be scored -1 and if the milestone has been maintained then it will be scored 0. These scores will then be summed to give a total score which could range between -6 (if a subject was able to perform all milestones at baseline but lost the ability and could not perform any milestones at Day 302) and +6 (if a subject was unable to achieve any milestones at baseline but could achieve all at Day 302).

The total number of new milestones achieved by each subject at Day 302 will be compared between treatment groups by ANCOVA with treatment group as the factor and subject age at screening, baseline Plasma NfL, and baseline number of milestones as covariates.

The following sensitivity analyses will be performed:

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

- The main analysis will be repeated using the PPS.
- The analysis will be repeated in the ITT, with an exception to assign the worst of last available observation and zero (0) new motor milestones achieved for subjects who terminate the study prior to the Day 302 assessment.

Biomarker

The baseline for NfL is defined as the last non missing assessment prior to the first dose of study treatment. For Plasma NfL, if the baseline is missing and a NF-L result is available post dose at Day 1 or Day 2 than the first available assessment post dose, will be used as baseline.

Values that are BLQ will be set to half of the diluted LLOQ (34.5 pg/mL NfL in CSF and 2.76 pg/mL for NfL in plasma).

Neurofilament light chain (NfL) concentration in CSF and in plasma will be analyzed.

The baseline concentrations and the change (including absolute and ratio to baseline) will be presented by visit.

The analysis is to compare the adjusted mean change from (ratio to) baseline to Day 302 between 50/28 mg Group and 12/12 mg Group using an ANCOVA model with treatment group as a factor and age at screening, baseline CSF/Plasma NFL and baseline HFMSE score as covariates.

Summary tables and LS-mean plots (±SE) for geometric mean ratio to baseline in NfL in CSF/plasma using MI data +ANCOVA.

Spaghetti plots by visit for NfL in CSF/plasma using observed data will be plotted for all subjects.

Analysis Methods for ACEND

Parents of subjects will complete the Assessment of Caregiver Experience with Neuromuscular Disease (ACEND) questionnaire at screening on Day 183 and Day 302. This assessment instrument has been designed to quantify the caregiver impact experienced by parents of children affected with severe neuromuscular diseases, including children with SMA (Matsumoto et al. 2011).

The ACEND includes a total of seven domains assessing physical impact (including feeding/grooming/dressing, sitting/play, transfers, and mobility) and general caregiver impact (including time, emotion, and finance) and each domain comprises several items. The total score for a domain with n items, each item assessed on ordinal scale from 1 to z, is derived as follows: 100 multiplied by (Mean of the n items in the domain -1) divided by (z-1). This total score will be on a scale of 0 to 100 with a higher score indicating a greater impact on the caregiver. At least two items for the time domain and one item for the remaining domains need to be non-missing for a total to be calculated; else the total score will be set to be missing.

The baseline for ACEND is defined as the last non missing assessment prior to the first dose of study treatment.

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

The main analysis is to compare the change from baseline for each domain at Day 302 between 50/28 mg Group and 12/12 mg Group using an ANCOVA model with treatment group as a factor and age at screening, baseline plasma NfL and baseline score as covariates. Missing data will be handled using the multiple imputation method.

Analysis Methods for PedsQL

Subjects will be evaluated using the Pediatric Quality of Life Inventory (PedsQL^{TM}) Measurement 4.0 Generic Core Scales and 3.0 Neuromuscular Module (Vami et al. 1999) at screening (the baseline assessment) , Day 183 and Day 302.

The questionnaires are specific to the age of the subject, and sites are instructed to get both subjects and caregivers to complete the same age specific questionnaire as was collected at baseline irrespective of whether or not the subjects cross an age boundary at a subsequent visit.

The PedsQL parent questionnaire is collected for children in the following age categories: 2-4, 5-7, 8-12. Four dimensions are collected: Physical, Emotional, Social and School functioning and each item is scored on a 5 point ordinal scale (0=Never, 1=Almost Never, 2=Sometimes, 3=Often, 4=Almost Always). The PedsQL patient questionnaire is collected for children in the following categories: 5-7, 8-12. Similar dimensions and 5-point ordinal scale are used as for the parents but for subjects aged 5-7 years a 3-point ordinal scale is collected, omitting the response levels of 1 and 3.

In the neuromuscular module, one parent questionnaire is collected for all subjects irrespective of age with three dimensions: “About my child’s neuromuscular disease”, “Communication” and “Family resources”. The same 5-point ordinal scale is collected for each question.

The patient neuromuscular disease questionnaire is collected for subjects in the following age categories: 5-7, 8-12. The questionnaire for subjects aged 5-7 years uses the 3-point ordinal scale as above and has only one dimension – “About my Neuromuscular disease”.

The questionnaires for the remaining two categories cover dimensions for “Communication” and “Family resources” and use the 5-point ordinal scale.

In scoring a dimension the first step is to reverse and linearly transform to a 0-100 scale (0=100, 1=75, 2=50, 3=25, 4=0), so higher score is indicative of a better health related quality of life. If greater than 50 percent of the items within a dimension are missing then the dimension score will not be computed, otherwise the mean score for the dimension will be calculated as the sum of items over the number of items answered.

A psychosocial health summary score, constructed from three dimensions, will be calculated as the sum of items over the number of items answered in the emotional, social and school functioning scales. A total score will be calculated as the sum of all the items over the number of items answered on all the scales. If greater than 50 percent of the items are missing, then the summary score or total score will be set to be missing.

For the neuromuscular module, a score for each dimension and then total score will be calculated in the same manner, no health summary scores are evaluated.

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

The baseline for PedsQL is defined as the last non missing assessment prior to the first dose of study treatment.

For the generic PedsQL and neuromuscular module, the change from baseline in the total score will be presented by visit and treatment group for both the parent and subject assessments. Similarly, the changes in the individual dimensions and health summary score will also be presented.

The change from baseline score (each total score and parent/subject evaluation separately) at Day 302 50/28 mg Group and 12/12 mg Group will be analyzed using ANCOVA model with treatment group as a factor and age at screening, baseline plasma NfL and baseline score as covariates. Missing data will be handled using the multiple imputation method. The imputation will be performed separately for each dimension and total and psychosocial summary scores. Due to the age specific nature of these questionnaires, subjects aged 2-4 years would not be expected to complete the self-evaluation and if this is the case then imputations will not be performed, and the subjects will be excluded from the analysis.

5.1.6. Number and Length of Hospitalizations

If the number of hospitalizations during the study is greater than 5 in total then an analysis will be performed as described in the following section using the rate at which they occur, otherwise just a listing will be provided.

For descriptive purposes, the aggregate hospitalization rate will be calculated for each group by dividing the total number of hospitalizations that occurred in the group by the total number of subject-years on study.

Annualized hospitalization rate will be calculated for each subject as the number of hospitalizations that the subject experienced divided by the number of days on study and this ratio multiplied by 365. Annualized hospitalization rate will be analyzed by a negative binomial regression model, adjusted for disease duration at Baseline. The logarithmic transformation of the time on the study will be included in the model as the “offset” parameter. If the data are under dispersed, or if the negative binomial regression model does not converge, a Poisson regression model with the same covariate will be used instead of the negative binomial regression model. Dispersion will be evaluated from the Pearson chi-squared statistic.

For each subject, the total time spent in hospital during the study will be calculated and prorated according to number of days on study. Based on the prorated time in hospital, treatment groups will be compared to an ANCOVA adjusting for each subject’s disease duration at Screening, age at symptom onset and baseline Plasma NfL.

5.1.7. Clinical Global Impression of Change (CGIC) [Physician, Caregiver] at Day 302

CGIC will be assessed at Days 29, 64, 183 and 302. The CGIC rating scale was developed as a brief standalone assessment of the clinician’s view of the patient’s global functioning after initiating a study medication [Guy 1976]. The CGIC provides an overall clinician-determined summary measure that takes into account all available information, including knowledge of the

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

patient's history, psychosocial circumstances, symptoms, behavior, and the impact of the symptoms on the patient's ability to function. The CGIC is administered by an experienced clinician who is familiar with the disease under study and the likely progression of treatment. The clinician makes a judgment about the total picture of the patient at each visit: the severity of the illness, the patient's level of distress and other aspects of impairment, and the impact of the illness on functioning. The CGIC is rated without regard to the clinician's belief that any clinical changes are or are not due to medication and without consideration of the etiology of the symptoms.

The CGIC is a 7-point scale that requires the clinician to assess how much the patient's illness has changed relative to a baseline state at the beginning of the intervention and is rated as follows: 1, very much improved; 2, much improved; 3, minimally improved; 4, no change; 5, minimally worse; 6, much worse; or 7, very much worse.

CGIC should be assessed consistently by the same rater for each study participant. A separate CGIC assessment will be performed by the Investigator (Principal Investigator or Sub investigator) and caregiver.

Clinical Global Impression of Change (CGIC) for the physician and caregiver at Day 302 will be analyzed as an ordered categorical score and as a dichotomized response: "much improved" or "very much improved" versus other responses.

The proportion of subjects with each CGIC score and with the dichotomized response of either "much improved" or "very much improved" will be summarized by visit and treatment group.

The difference in CGIC dichotomized response rate between 50/28 mg Group and Control Group will be analyzed using a logistic regression model with terms for treatment group and disease duration. Participants who die or withdraw from the study, prior Day 302, will be counted as 'non responders'.

The odds ratio, its 95% CI and the p-value for 50/28 mg Group compared to the Control Group will be presented. The proportion of CGIC responders at other visits will also be analyzed in a similar way. In a situation where the model does not converge then exact logistic regression (Hirji, K. F. 2005) will be performed.

5.1.8. Number of Serious Respiratory Events

All adverse events that are coded into the System Organ Class (SOC) of respiratory, thoracic, and mediastinal disorders, either as their primary SOC will be considered respiratory events. Only those classified as serious will be analyzed. By utilizing the above approach, rare occurrence of serious respiratory events in the Infection and Infestation SOC and Investigation SOC (if any) may not be included. However, a detailed review of the AE listing will be conducted to identify such events.

The number of serious respiratory events during the study will be analyzed using the rate at which they occur. For descriptive purposes, the aggregate rate will be calculated for each group by dividing the total number of serious respiratory events that occurred in the group by the total number of subject-years on study.

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

Annualized serious respiratory event rate will be calculated for each subject as the number of serious respiratory events that the subject experienced divided by the number of days on study and this ratio multiplied by 365.

No formal statistical comparison is planned.

5.1.9. Ventilator Use

The participant's ventilator use and respiratory aids will be collected at every study visit. If ventilation is used daily, the average number of hours per day for the past 7 days will be recorded.

The ventilation use will be summarized by visit and listed.

5.1.10. Change in the Parent Assessment of Swallowing Ability Scale (PASA)

Parent Assessment of Swallowing Ability (PASA) will be assessed using the PASA questionnaire. Caregivers will be asked a series of questions regarding the mealtime behavior of the participant.

The PASA questionnaire was developed by a Biogen team in order to assess the signs and symptoms of dysphagia. This questionnaire consists of 33 items across 4 domains that cover general feeding, drinking liquids, eating solid foods, and assessment of swallowing concerns. The first 3 of these domains are generally assessed with 5 levels of response (Never, Rarely, Sometimes, Often, and Always), although 2 items are assessed with a simple "Yes"/"No" answer. In the final domain, the assessment of swallowing concerns has 4 levels of response: Strongly Agree, Agree, Disagree, and Strongly Disagree. In answering each item, the caregiver is directed to consider the previous 7 days.

The number, mean score, and percentage of subjects scoring each category in the PASA Section 1 – General Feeding will be presented by visit. In addition, a stacked bar chart will be used to summarize the proportion of subjects in each category over time for each item. Depending on the number of subjects who complete PASA Sections 1,2, 3, and 4, similar presentations may also be made or alternatively the results from these visits will be listed.

[redacted]

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

5.1.12. Subgroup analyses

Summary statistics by visit for the endpoints of HFMSE, RULM and number of new WHO milestones will be presented by age category and by region. It is not planned to do ANCOVA analysis due to the small patient numbers.

[redacted]

- [redacted]
- [redacted]

[redacted]

5.2. Comparisons to external data

The following table summarizes the endpoints, group compared, timepoint and analysis method where additional supplementary comparisons to the DEVOTE higher dose group will be performed:

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

Time points	Matched group compared	Analysis
Plasma NfL		
Plasma NfL at Month 6	PS matched CS4 Sham	MI +ANCOVA [Age at first dose, Baseline Plasma NfL, Baseline HFMSE]
	PS matched nusinersen - CS4 only	MI +ANCOVA [Age at first dose, Baseline Plasma NfL, Baseline HFMSE]
HFMSE		
Change to Month 9 [Repeated for Month 6]	PS matched CS4 Sham	MI +ANCOVA [Age at first dose, Baseline Plasma NfL, Baseline HFMSE]
Change to Month 9	PS matched nusinersen – CS4+CS11	MI +ANCOVA [Age at first dose, Baseline HFMSE]
Change to Month 9 [Repeated for Month 6]	PS matched nusinersen – CS4 only	MI +ANCOVA [Age at first dose, Baseline Plasma N-FL, Baseline HFMSE]
Change to Month 9	PS with caliper matched nusinersen A (Clinical trials and PMS)	MI +ANCOVA [Age at first dose, Baseline HFMSE]
Change to Month 9	PS with caliper matched nusinersen B (Clinical trials)	MI +ANCOVA [Age at first dose, Baseline HFMSE]
Change to Month 9 [Repeated for Month 6]	PS with caliper matched CS4 Sham	MI +ANCOVA [Age at first dose, Baseline Plasma NfL, Baseline HFMSE]
Proportion of subjects achieving: >=3 point improvement from baseline at Month 9[Repeated for Month 6]	PS matched CS4 Sham	MI + logistic regression [Age at first dose, Baseline Plasma NfL, Baseline HFMSE]
Proportion of subjects achieving: >=3 point improvement from baseline at Month 9	PS matched nusinersen – CS4+CS11	MI + logistic regression [Age at first dose, Baseline HFMSE]
Proportion of subjects	PS matched	MI + logistic regression [Age at first dose,

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

Time points	Matched group compared	Analysis
achieving: >=3 point improvement from baseline at Month 9[Repeated for Month 6]	nusinersen – CS4 only	Baseline Plasma NfL and Baseline HFMSE]
Proportion of subjects achieving: >=3 point improvement from baseline at Month 9	PS with caliper matched nusinersen A (Clinical trials and PMS)	MI + logistic regression [Age at first dose, Baseline HFMSE]
Proportion of subjects achieving: >=3 point improvement from baseline at Month 9	PS with caliper matched nusinersen B (Clinical trials)	MI + logistic regression [Age at first dose, Baseline HFMSE]
Proportion of subjects achieving: >=3 point improvement from baseline at Month 9[Repeated for Month 6]	PS with caliper matched CS4 Sham	MI + logistic regression [Age at first dose, Baseline Plasma NfL, Baseline HFMSE]
RULM		
Change to Month 9 [Repeated for Month 6]	PS matched CS4 Sham	MI +ANCOVA [Age at first dose, Baseline Plasma NfL, Baseline RULM]
Change to Month 9	PS matched nusinersen – CS4+CS11	MI +ANCOVA [Age at first dose, Baseline RULM]
Change to Month 9 [Repeated for Month 6]	PS matched nusinersen – CS4 only	MI +ANCOVA [Age at first dose, Baseline Plasma NfL, Baseline RULM]
Change to Month 9	PS with caliper matched nusinersen B (Clinical trials)	MI +ANCOVA [Age at first dose, Baseline RULM]
Change to Month 9 [Repeated for Month 6]	PS with caliper matched CS4 Sham	MI +ANCOVA [Age at first dose, Baseline Plasma NfL, Baseline RULM]
Total number of new WHO motor milestones		

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

Time points	Matched group compared	Analysis
Change to Month 9 [Repeated for Month 6]	PS matched CS4 Sham	ANCOVA [Age at first dose, Baseline plasma NfL, Baseline number of WHO milestones]
Change to Month 9	PS matched nusinersen – CS4+CS11	ANCOVA [Age at first dose, Baseline number of WHO milestones]
Change to Month 9 [Repeated for Month 6]	PS matched nusinersen – CS4 only	ANCOVA [Age at first dose, Baseline plasma NfL, Baseline number of WHO milestones
Change to Month 9	PS with caliper matched nusinersen B (Clinical trials)	ANCOVA [Age at first dose, Baseline number of WHO milestones]
Change to Month 9 [Repeated for Month 6]	PS with caliper matched CS4 Sham	ANCOVA [Age at first dose, Baseline number of WHO milestones]

Square brackets in the 3rd column denotes the covariates to include in the ANCOVA/Logistic regression model. If analysis is also using multiple imputation, then these covariates will be included in the imputation step.

6. Safety Endpoints

The analysis of safety will be performed using the Safety Set. The main analysis will be to compare the two arms of 232SM203. Presentations will also be included for AEs, SAEs and related AEs comparing sham from the matched CS4 Sham to the 232SM203 higher dose arm.

AEs will be analyzed based on the principle of treatment emergence where treatment emergence will be relative to the first dose of nusinersen. An AE will be regarded as treatment-emergent if it was present prior to receiving the first dose of nusinersen and subsequently worsened in severity or was not present prior to receiving the first dose of nusinersen but subsequently appear. Refer to Appendix A for more details about AE with missing start or stop dates.

A participant having the same AE preferred term more than once will be counted only once in the incidence for that event. Multiple occurrences of the same adverse event for the same subject will all be counted in the frequency for that adverse event.

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

The occurrence of the AE with the greatest severity will be used in the calculation of incidence by severity. The AE relationship to treatment or LP/sham procedure is captured as related or not related. In the event the relationship is missing then this will be considered related.

Due to the long half-life of nusinersen, analyses of treatment-emergent adverse events will include all events reported during the study. Adverse Events will be coded using the Medical Dictionary for Regulatory Activities.

The following presentations will be shown:

- an overall summary showing, for each treatment group, the number and percentage of subjects with an adverse event, a mild, moderate or severe event, an event related to study drug, an event related to lumbar puncture /sham procedure, a serious event, a serious event related to study drug, an event leading to study drug withdrawal, an event leading to study withdrawal, and a fatal event.
- incidence and frequency by system organ class and preferred term
- incidence and frequency by preferred term
- incidence and frequency, by preferred term, in at least 2 subjects in any treatment group
- incidence of adverse events by maximum severity by preferred term
- incidence and frequency of severe events by system organ class and preferred term
- incidence and frequency of events related to study drug by system organ class and preferred term
- incidence and frequency of events related to lumbar puncture /sham procedure (as assessed by the investigator) by system organ class and preferred term
- incidence by system organ class and preferred term that occurred within 24 hours of Dosing or Sham Procedure
- incidence by system organ class and preferred term that occurred within 72 hours of Dosing or Sham Procedure
- incidence and frequency of serious adverse events by system organ class and preferred term
- incidence and frequency of serious adverse events related to study drug by system organ class and preferred term
- incidence of events that led to fatal outcome
- incidence of events leading to study drug withdrawal by system organ class and preferred term

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

- incidence of events leading to withdrawal from study by system organ class and preferred term
- incidence and frequency of adverse events over time by system organ class and preferred term.
- incidence and frequency of adverse events antibody status by system organ class and preferred term.

A separate table will also be presented to show an overall summary of COVID-19 pandemic related AEs.

Listings of the following events will be produced.

- AEs
- SAEs
- AE with fatal outcome
- AEs led to study drug withdrawal
- AEs led to withdrawal from study
- AEs related to lumbar puncture/sham procedure
- AEs related to study drug

Follow-up adjusted incidence rate

Follow-up adjusted incidence rate will be summarized by system organ class and preferred term by study group. Follow-up adjusted incidence rate, defined as the number of subjects who experienced an event divided by the total of entire follow-up time among the subjects in the analysis population. The entire follow-up time for subject is defined as the time from the first dose until the end of study day. Each subject will be counted only once within each category. A similar table will also be based on total number of events divided by the total of entire follow-up time among the subjects in the analysis population. This may count a subject more than once within each category if they experienced an event more than once.

Note on capture of severity and seriousness in CS4

In CS4 the eCRF captured detail for each investigator term for decreases in severity and changes from serious to non-serious. In 232SM203, the decrease in severity and change from serious to non-serious is not captured. In all studies worsening in severity and moving from non-serious to serious is captured.

In presenting CS4 as part of 232SM203 reporting, consistent with how these studies were originally reported in the following two scenarios:

1) where the severity decreases

2) where a serious event becomes non-serious

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

Only the first record will be considered treatment emergent so the actual reporting approach will be consistent with 232SM203.

6.1.1. Concomitant therapy and Procedure

A concomitant therapy is any drug or substance administered between screening and the final study visit/telephone call. Participants are instructed to contact their Investigators before taking any new medications, including nonprescription drugs and herbal preparations.

Throughout the study, Investigators or designated licensed physicians involved in the study may prescribe concomitant therapies (procedures) or treatments deemed necessary for AEs or to provide adequate supportive care.

Participants are prohibited from receiving other experimental or approved agents for the

treatment of SMA, including gene therapy, biological agent, or device, during the study. This includes marketed agents at experimental doses that are being tested for the treatment of SMA (oral albuterol/salbutamol, riluzole, carnitine, sodium phenylbutyrate, valproate, hydroxyurea, etc.).

All concomitant therapies will be coded using the World Health Organization drug

dictionary (WHO Drug). A concomitant procedure is any therapeutic intervention (e.g., surgery/biopsy, physical therapy) or diagnostic assessment (e.g., blood gas measurement, bacterial cultures) performed between Screening and the final study visit/telephone call. Concomitant procedures will be coded using MedDRA.

For the purposes of analysis, a concomitant therapy (including medication or procedure) is defined as any therapy that was taken or administered on or after the first injection of nusinersen. (appendix A). This includes therapies and procedures that were started prior to the initiation of injection of nusinersen if their use continued on or after the first injection of nusinersen. The number and percentage of subjects who were taking each type of concomitant therapy and procedures at baseline and during the study will be presented by preferred term.

The number and percentage of subjects who underwent each type of concomitant procedure during the study will be presented.

Concomitant therapies related to LP/Sham procedure will be summarized separately.

If there are a sufficient number of tests or treatments reported with indication of COVID-19, a separate summary of these will be provided.

Concomitant medications will also be reviewed by a medical reviewer prior to the lock to determine disallowed medications according to the protocol. These will be summarized separately.

6.1.2. Clinical laboratory data

Laboratory data will be evaluated to determine the incidence of abnormalities that emerge during the course of the study. Changes in laboratory evaluations will be presented relative to baseline, which is defined as the closest visit prior to the first dose.

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

The following clinical laboratory parameters are to be assessed:

- Hematology panel: complete blood count with differential and platelet count (hematocrit, hemoglobin, platelets, red blood cell count [RBC], white blood cell count [WBC], basophils, eosinophils, lymphocytes, monocytes, neutrophils)
- Blood chemistry panel: albumin, total bilirubin, direct and indirect bilirubin, alkaline phosphatase, alanine aminotransferase (ALT) (SGPT), aspartate aminotransferase (AST) (SGOT), gamma-glutamyl transferase (GGT), sodium, potassium, calcium, chloride, phosphate, blood urea nitrogen (BUN), creatinine, cystatin, creatine phosphokinase, creatine kinase, bicarbonate (CO2), glucose, total protein.
- Urinalysis: specific gravity, pH, protein, glucose, ketones, bilirubin, blood, red blood cells, white blood cells, epithelial cells, bacteria, casts and crystals
- Urine total protein assessed by local laboratories CSF analysis: RBC, WBC, protein, glucose.
- Coagulation: PT (prothrombin time), aPTT (activated partial thromboplastin time), and INR (international normalized ratio).

Each hematology, blood chemistry, coagulation and CSF laboratory parameter will be flagged as “low”, “normal” or “high” relative to the parameter’s normal range or as “unknown” if no result is available.

For each urinalysis laboratory parameter, the number and percentage of subjects experiencing postdose shifts to abnormal will be summarized.

For each hematology, blood chemistry, coagulation and CSF parameter, the number and percentage of subjects experiencing postdose shifts to ‘low’ or ‘high’ will be summarized. In each summary, the denominator for the percentage is the number of subjects at risk for the shift. The number at risk for the shift to low is the number of subjects whose baseline value was not low and who had at least one post baseline value. The number at risk for the shift to high is the number of subjects whose baseline value was not high and who had at least one-post baseline value. Subjects will be counted only once for each parameter and each shift regardless of how many postdosing assessments had that type of shift. All post baseline data will be used in the shift tables, regardless of whether a scheduled or unscheduled visit.

Also, for each parameter, the incidence of shift to low will be summarized using the minimum post baseline values. Shift to low includes subjects with a normal, high, or unknown baseline value and at least one post baseline value of the given test. Similarly, the incidence of shift to high will be summarized using the maximum post baseline values. Shift to high includes subjects with a low, normal, or unknown baseline value and at least one post baseline value.

To evaluate potential serious hepatotoxicity subjects with a post baseline AST and/or ALT value ≥3 times the upper limit of normal (ULN) and a post baseline bilirubin value >2 times ULN at any time, not necessarily concurrent, will be listed with their values. In addition, a plot will be presented with each subject’s maximum post baseline AST or ALT value relative to the ULN

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

against the subject's maximum post baseline bilirubin value relative to the ULN; values do not have to be concurrent. All post baseline data will be used to identify abnormalities, regardless of whether a scheduled or unscheduled visit.

Summary statistics for actual values and change from baseline in laboratory values will be summarized by treatment group and visit. Box plots for chemistry, hematology and CSF showing mean value for each treatment group at each visit will also be presented.

Due to the impact of the COVID-19 pandemic on transport infrastructure, public health or humanitarian emergencies, certain sites may only be able to test blood and urine samples at a local laboratory instead of central one. The use of local labs which will use different techniques/equipment and reference ranges may introduce variability in the results.

To assess the variability of the results two summaries will be presented for actual and change from baseline by visit:

- summarize local laboratory and central laboratory results together
- summarize only central laboratory results

Listings of all chemistry, hematology, coagulation, CSF, urinalysis values and pregnancy test (serum and urine) will be provided.

For platelets a summary of the proportion of participants with a post baseline platelet count below the lower limit of normal on at least 2 consecutive measurements will be presented.

For study 232SM203, the urine protein positive (trace, +1, +2, +3, +4 or >0.2g/L) results will be presented, number of subjects with positive result at baseline, with a negative result at baseline and at least one and at least two positive post baseline result respectively.

6.1.3. Neurological Examinations

Neurological examinations include assessment of mental status, level of consciousness, cranial nerves, reflexes, motor system, coordination/cerebellar function and sensation – temperature and vibration. These are to be assessed at Screening, predose and at 3, 6, and 24 hours postdose on dosing Days 1, 15, predose and at 1 and 3 hours postdose on Days 29, 64, 135, 183, and 279; and at Day 302/ET.

The results collected for the majority of the tests are classified as 'normal' or 'abnormal', however the assessment of sensations is reported as 'present' or 'absent' and the assessment of reflexes is captured on an ordinal scale. For each test it is recorded if secondary to SMA.

Acute Effects After Dosing: **Reflex assessments**

On each dosing day, each predose reflex will be compared with all the post dosing assessments with 24 hours to see if the subject has changed (gained or lost) or maintained, reflexes. For this comparison the predose reflex status will be categorized into two categories: 0 and >=1. For each

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

reflex, visit, predose category the number and proportion of subjects who change or maintained will be presented.

Acute Effects After Dosing: Other assessments with binary response (Normal/Abnormal)

On each dosing day, the shifts from predose assessment on that day to postdose (i.e., from predose to 3, 6 and 24 hours postdose) will be determined. For each test and post dosing time point, the number and proportion of subjects who moved from 'Normal' to 'Abnormal' will be presented, for sensations the number and proportion who move from 'Present' to 'Absent' will be presented.

Chronic effects: Reflex assessments

For each reflex the baseline reflex will be compared (the predose value on Day 1) to the pre dose value on later dosing days (Day 1, 15, 29, 64, 135,183, 279 and on Day 302/ET) to see if the subject has changed (gained or lost) or maintained, reflexes. For this comparison the predose reflex status will be categorized into two categories: 0 and >=1. For each reflex, visit, predose category the number and proportion of subjects who change or maintained will be presented.

Chronic effects: Other assessments with binary response (Normal/Abnormal)

For each test the shift from baseline to (the predose value on Day 1) the predose value on later dosing days (Day 1, 15, 29, 64, 135,183, 279 and on Day 302/ET)to determine, the number and proportion of subjects who moved from 'Normal' to 'Abnormal' will be presented, for sensations the number and proportion who move from 'Present' to 'Absent' will be presented.

Figures of subjects neurological exams will also presented.

6.1.4. ECG

ECGs are to be recorded at Days 1, 15, 29, 64, 135, 183, 279 and 302/ET, predose and post dosing at various time points on dosing days and end predose at 302/ET. The ECGs are assessed at a central reading laboratory and the results provided as external vendor data. On the eCRF the investigator interpretation is collected as normal, abnormal and abnormal AE and the investigator determination will be used in analyses.

The ECG test includes heart rate, QT interval, QTcF interval and RR interval. Summary statistics for actual values and change from baseline in each ECG parameter will be presented by treatment group and visit.

ECGs will be analyzed using two approaches.

Qualitative analysis

The determination of whether an ECG is abnormal not AE, abnormal AE made by the Investigator will be used for this qualitative analysis.

The number and percentage of subjects with shifts from normal to each of the categorical values denoting an abnormal scan (abnormal not AE, abnormal AE) will be summarized by treatment group. All post baseline data will be used to identify abnormalities, regardless of whether a scheduled or unscheduled visit.

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

Corrected Fridericia QT interval QTcF abnormalities (outliers)

QTcF will be examined to determine the incidence of clinically relevant abnormalities. The number of subjects evaluated and the number of subjects with clinically relevant abnormalities will be presented.

The first criteria for clinically relevant post dosing abnormalities (secondary safety endpoint) that will be considered is:

- Post baseline QTcF of > 500 msec and Maximum increase from baseline to post baseline QTcF >60 msec

Other criteria are:

- Maximum increase from baseline QTcF > 30 to 60 msec
- Maximum increase from baseline QTcF > 60 msec
- Maximum post baseline QTcF > 480 to 500 msec
- Maximum post baseline QTcF > 500 msec

ECG summaries will be presented for all subjects and for subjects with post baseline. Listings will include both the central and eCRF information.

6.1.5. Vital Signs

The analysis of vital signs will be approached in two ways.

Vital signs are to be measured at Screening, predose and at 1, 2, 4, 6, and 24 hours postdose on dosing Days 1, 15, predose at 1 and 6 hours postdose on Days 29, 64, 135, 183, and 279 and on Day 302/ET At each of these times, temperature, heart rate, respiratory rate, systolic and diastolic blood pressure, and pulse oximetry will be measured.

Acute Effects After Dosing

On each dosing day, the change in each vital sign from predose on that day to postdose (i.e., from predose to 1 hour postdose, to 2 hours postdose, etc.) will be calculated. Summaries of actual values and change from predose for each dosing day will be presented.

Chronic Effects

To examine for possible chronic effects, the change from baseline (the predose value on Day 1) to the predose value on later dosing days and to Day 302/ET will be determined. Summaries of actual values and change from baseline for subsequent dosing days (Days 15, 29, 64,135, 183, and 279) and for Day 302/ET will be presented.

Vital signs (temperature, pulse, systolic and diastolic blood pressure and pulse oximetry) will also be examined to determine the incidence of potentially clinically relevant abnormalities.

Shift analysis

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

The number of subjects evaluated and the number of subjects with potentially clinically relevant abnormalities will be presented. All post baseline data will be used to identify abnormalities, regardless of whether a scheduled or unscheduled visit. The criteria for potentially clinically relevant postdose abnormalities are shown in Table 6 below:

Table 6: Criteria to determine potentially clinically relevant abnormalities in vital signs.

Vital Sign	Criteria for Abnormalities
Temperature	<36°C
	>38°C
Pulse	<60 bpm
	>100 bpm
Systolic Blood Pressure	<90 mmHg
	>140 mmHg
	>160 mmHg
Diastolic Blood Pressure	<50 mmHg
	>90 mmHg
	>100 mmHg
Weight	7% or more decrease from baseline
	7% or more increase from baseline
Respiratory Rate	<12 breaths/min
	>20 breaths/min

6.1.6. Growth parameters

Growth parameters include weight for age, and ulnar length.

The growth parameters are to be assessed during screening, on dosing days (Days 1, 15, 29, 64, 135, 183 and 279) and on Day 302/ET.

The 2000 CDC Growth Charts (ages 2 to <20 years) will be used to assess the weight change for older subjects. The National Center for Health Statistics provides a SAS macro which can be downloaded from their website [https://www.cdc.gov/growthcharts/clinicalcharts.htm] and this will be utilized to calculate the weight for age percentiles for each participant in the later onset population.

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

Summary statistics for actual values and change from baseline in each growth parameter will be presented by treatment group and visit.

Baseline growth parameters will be summarized and also the frequency and percentage for weight and height for age will also be presented for each of in the following categories: ≤ 1st, ≤ 3rd, ≤5th, ≤15th, ≤25th, ≤50th and > 50th percentile.

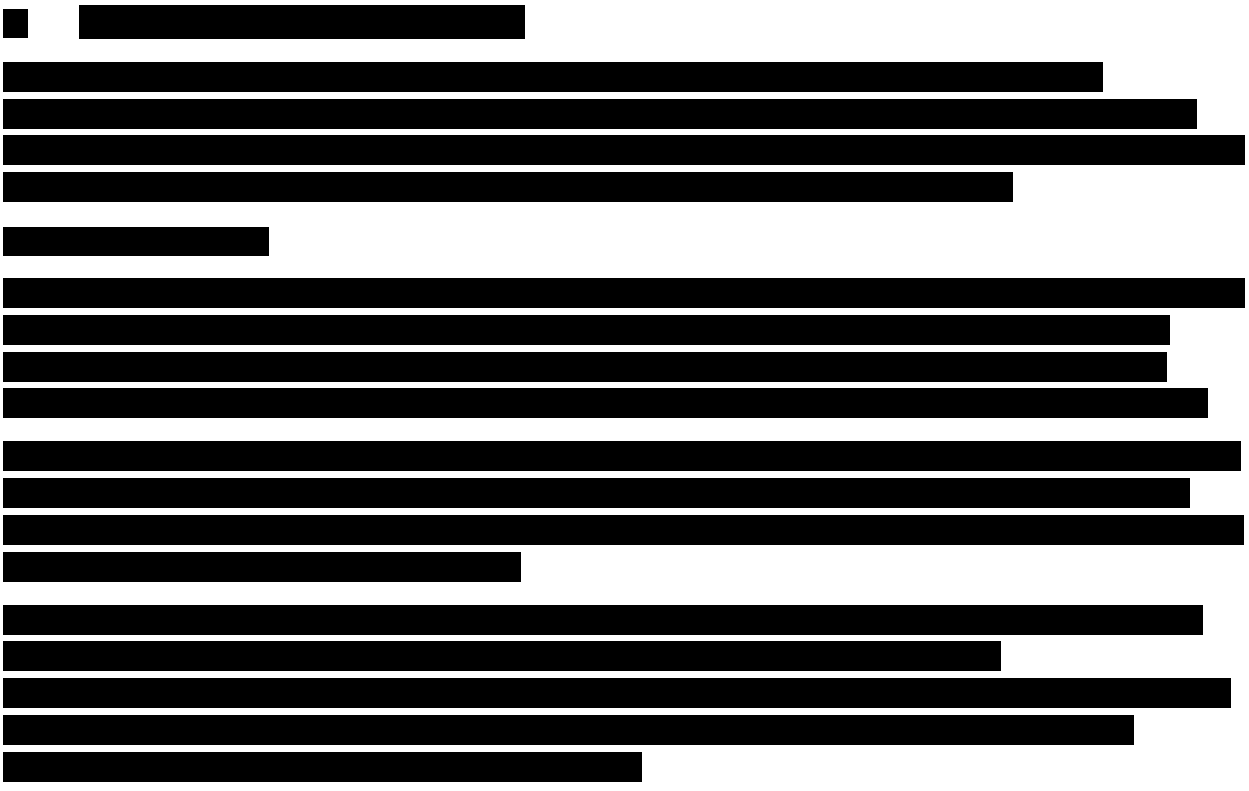
6.1.7. LP opening pressure

LP opening pressure will only be measured at Days 1, 15, and 279 to avoid potential unblinding. A lack of published data exists in terms of the variability of opening pressure between and within subjects. It is suspected that a simple change in position of a subject during the measurement or change of instrument used contributes to variability in the measurement.

The actual value and change from baseline to each visit will be summarized by treatment groups all the data collected will be listed.

6.1.8. X-ray Examination of Spine

The X-rays will be used to determine the severity of scoliosis by measuring the Cobb angle in the later-onset SMA subjects. The spine X-ray is performed at screening with the participant in a sitting or supported sitting position following an image acquisition guideline. Summary and listing of baseline Cobb angle will be presented.



DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

[redacted]

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

- [redacted]
- [redacted]

[redacted]

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

[redacted]

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

9. Pharmacodynamic Endpoints

As detailed earlier NfL is defined as a secondary endpoint.

[redacted] In addition, plasma and CSF PK concentrations will be analyzed with plasma and CSF biomarkers to establish PK/PD correlation. PK/PD and exposure-response may also be conducted on efficacy. The results of this study may be combined with other nusinersen studies to perform population PK/PD analyses, including exposure-response analysis.

10. Other Analyses

Not applicable.

11. Statistical Considerations for Interim Analysis

Not applicable

12. Changes from Protocol-Specified Analyses

Added NF-L as a secondary endpoint.

13. Summary of Changes from the Previous Version of the SAP

14. References

[redacted]

Arkblad E, Tulinius M, Kroksmark AK, et al. A population-based study of genotypic and phenotypic variability in children with spinal muscular atrophy. Acta Paediatr. 2009;98(5):865-72.

[redacted]

Berry JD, Miller R, Moore DH, et al. The Combined Assessment of Function and Survival (CAFS): a new endpoint for ALS clinical trials. Amyotroph Lateral Scler Frontotemporal Degener. 2013;14(3):162-8. Epub 2013/01/17.

Bohdana Ratitch, Ilya Lipkovich, Michael O'Kelly. Combining Analysis Results from Multiply Imputed Categorical Data. PharmaSUG 2013 - Paper SP03.

Coovert DD, Le TT, McAndrew PE, et al. The survival motor neuron protein in spinal muscular atrophy. Hum Mol Genet. 1997;6(8):1205-14.

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

De Sanctis R, Coratti G, Pasternak A, et al. Developmental milestones in type I spinal muscular atrophy. Neuromuscul Disord. 2016;26(11):754-759. Epub 2016/10/05.

Elizabeth A. Stuart. "Matching Methods for Causal Inference: A Review and a Look Forward." Statistical Science, 25(1) 1-21 February 2010

Feldkötter M, Schwarzer V, Wirth R, et al. Quantitative analyses of SMN1 and SMN2 based on real-time lightCycler PCR: fast and highly reliable carrier testing and prediction of severity of spinal muscular atrophy. Am J Hum Genet. 2002;70(2):358-68.

Finkel RS, Mercuri E, Meyer OH, et al. Diagnosis and management of spinal muscular atrophy: Part 2: Pulmonary and acute care; medications, supplements and immunizations; other organ systems; and ethics. Neuromuscul Disord. 2017 Epub 2017/11/23.

Finkel RS, Mercuri E, Meyer OH, et al. Diagnosis and management of spinal muscular atrophy: Part 2: Pulmonary and acute care; medications, supplements and immunizations; other organ systems; and ethics. Neuromuscul Disord. 2018;28(3):197-207. Epub 2017/11/23.

Finkelstein DM, Schoenfeld DA. Combining mortality and longitudinal measures in clinical trials. Stat Med 1999; 18 (11):1341-54.f

Glanzman AM, Mazzone E, Main M, et al. The Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders (CHOP INTEND): test development and reliability. Neuromuscul Disord. 2010;20(3):155-61.

Glanzman AM, McDermott MP, Montes J, et al. Validation of the Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders (CHOP INTEND). Pediatr Phys Ther. 2011;23(4):322-6.

Guy W. ECDEU Assessment Manual for Psychopharmacology. Rockville, MD: U.S. Department of Health, Education, and Welfare; 1976.

Haataja L, Mercuri E, Regev R, et al. Optimality score for the neurologic examination of the infant at 12 and 18 months of age. J Pediatr. 1999;135(2 Pt 1):153-61.

Hirji, K. F. (2005). Exact Analysis of Discrete Data. Boca Raton: Chapman and Hall.

ICH E9 (R1) Addendum 2017

Iannaccone ST, Hynan LS, Morton A, et al. The PedsQL in pediatric patients with Spinal Muscular Atrophy: feasibility, reliability, and validity of the Pediatric Quality of Life Inventory Generic Core Scales and Neuromuscular Module. Neuromuscul Disord. 2009;19(12):805-12.

[redacted]

[redacted]

[redacted]

Jedrzejowska M, Milewski M, Zimowski J, et al. Incidence of spinal muscular atrophy in Poland--more frequent than predicted? Neuroepidemiology. 2010;34(3):152-7.

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

Lefebvre S, Burlet P, Liu Q, et al. Correlation between severity and SMN protein level in spinal muscular atrophy. Nat Genet. 1997;16(3):265-9.

Little RJA, Rubin DB. Statistical Analysis with Missing Data. 2 nd ed. New York: John Wiley & Sons, Inc; 2002.

Main M, Kairon H, Mercuri E, et al. The Hammersmith functional motor scale for children with spinal muscular atrophy: a scale to test ability and monitor progress in children with limited ambulation. Eur J Paediatr Neurol. 2003;7(4):155-9.

[redacted]

Matsumoto H, Clayton-Krasinski DA, Klinge SA, et al. Development and initial validation of the assessment of caregiver experience with neuromuscular disease. J Pediatr Orthop. 2011;31(3):284-92.

Mazzone ES, Mayhew A, Montes J, et al. Revised upper limb module for spinal muscular atrophy: Development of a new module. Muscle Nerve. 2016 Epub 2016/10/04.

[redacted]

Mercuri E, Finkel RS, Muntoni F, et al. Diagnosis and management of spinal muscular atrophy: Part 1: Recommendations for diagnosis, rehabilitation, orthopedic and nutritional care. Neuromuscul Disord. 2018;28(2):103-115. Epub 2017/11/23.

[redacted]

O'Hagen JM, Glanzman AM, McDermott MP, et al. An expanded version of the Hammersmith Functional Motor Scale for SMA II and III patients. Neuromuscul Disord. 2007;17(9-10):693-7.

Permutt T, Li F. Trimmed means for symptom trials with dropouts. Pharm Stat. 2017;16(1):20-28. Epub 2016/08/15.

Prior TW, Swoboda KJ, Scott HD, et al. Homozygous SMN1 deletions in unaffected family members and modification of the phenotype by SMN2. Am J Med Genet A. 2004;130A(3):307-10.

Rosenbaum 2010, Design of Observational Studies

[redacted]

Santner TJ, Snell MK. Small-Sample Confidence Intervals for p 1 - p 2 and p 1 /p 2 in 2 × 2 Contingency Tables. J Am Statist Assoc. 1980;75:386-94.

Schafer JL. Analysis of Incomplete Multivariate Data. 1 ed. London: Chapman & Hall/CRC Press; 1997. 430 p.

Schafer JL. Multiple imputation: a primer. Stat Methods Med Res. 1999;8(1):3-15

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

[redacted]

Varni JW, Seid M, Rode CA. The PedsQL: measurement model for the pediatric quality of life inventory. Med Care. 1999;37(2):126-39.

WHO Multicentre Growth Reference Study Group. Assessment of differences in linear growth among populations in the WHO Multicentre Growth Reference Study. Acta Paediatr Suppl. 2006;450:56-65.

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

Wirth B, Brichta L, Schrank B, et al. Mildly affected patients with spinal muscular atrophy are partially protected by an increased SMN2 copy number. Hum Genet. 2006;119(4):422-8.

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

15. APPENDICES

APPENDIX A.

i Seeds for MI

Table 9: Seeds used for each endpoint in MI for Later-Onset SMA population

Endpoint	Seed
HFMSE	25980495
RULM	47886416
NfL in plasma	7953683
NfL in CSF	4711630
PedsQL	12676424
ACEND	68757534

ii Adverse Events

In the situation where worsening in severity (including seriousness) occurs for an adverse event, study sites are instructed to enter an end date and start a new record for the adverse event. In the new record, the changed severity is to be recorded and the start date will be set to be the end date of the previous record. Data linking those records are collected in the data base. Consider three scenarios:

- The first AE record occurs prior to the first dose, the second AE record occurs on or after the first dose, and the severity increases: Only the second record will be counted as treatment-emergent.
- Both records occur on or after the first dose: If the AE severity on the second record is worse than the severity on the first record, then count both records as treatment-emergent.
- Both records occur on or after the first dose: If the AE one is severe and the other serious then count both records as treatment-emergent.

Of note, when counting the total number of treatment-emergent events, events linked through change in severity will still be counted as separate events.

For events with missing start or stop dates, the following criteria will be used for the purpose of identifying treatment-emergent adverse events:

- If both the start and stop date for a particular event are missing, then the event is considered to have occurred on or after the first dose of study treatment.

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

- If the start date for a particular event is missing and the stop date/time falls after the first dose date/time, then the event is considered to have occurred on or after the first dose of study treatment.
- If the start time is missing and the start date is same as the first dosing date, then the event is considered to have occurred on or after the first dose of study treatment.
- If it cannot be determined whether or not an event has occurred on or after dosing due to a missing or partial date, then the event will be assumed to have occurred on or after the first dose for the purpose of identifying treatment-emergent adverse events.

Specifically, let AESTDT denote the start date of an adverse event and TRTSTDT be the start date of study treatment. For the purpose of identifying the treatment emergent adverse events, the following algorithm will be used for the imputation of missing or partial date:

- If AESTDT is completely missing or the year is missing, then impute AESTDT to TRTSTDT.
- If, in AESTDT, year is present and month/day are missing and year is equal to the year portion of TRTSTDT, then impute the month/day portion of AESTDT to the month/day portion of TRTSTDT.
- If, in AESTDT, year is present and month/day are missing and year is not equal to the year portion of TRTSTDT, then impute the month/day portion of AESTDT to January 01.
- Consider the situation in AESTDT where year and month are present with only day missing. If the year and month are the same as those for TRTSTDT, then impute day in AESTDT with day in TRTSTDT. Otherwise, impute the day in AESTDT with the first day of the month.

It is important to emphasize that the imputed date will not be used for calculations such as onset and duration of an adverse event.

iii Concomitant Therapy and Procedure

In order to define concomitant therapies with missing start or stop dates, the following additional criteria will be used:

- If both the start and stop dates of a particular therapy were missing, that therapy is considered concomitant.
- If the start date of a therapy was missing and the stop date of that therapy fell on or after the date of dosing, that therapy is considered concomitant.
- If the start date of a therapy was prior to the date of dosing and the stop date of that therapy was missing and the therapy was listed as continuing, that therapy is considered concomitant.

DocuSign Envelope ID: 666318C9-A78D-44C1-96B3-335ADB08F090

- If the start date of a therapy was prior to the date of dosing and the stop date of that therapy was missing and the therapy was listed as *not* continuing, that therapy is considered concomitant.
- If the start/stop date of a therapy is partial then where it is not possible to rule out that it was not taken concomitantly it will be considered concomitant.

Denote the end date of medications as CMENDT and the study treatment start date as TRTSTDT. The medication is classified concomitant provided any of the following is NOT true:

- CMENDT is complete and CMENDT is less than TRTSTDT.
- Day of CMENDT missing and year/month of CMENDT is strictly before year-month of TRTSTDT.
- Month of CMENDT is missing and year of CMENDT is strictly before year of TRTSTDT.

DocuSign

Certificate Of Completion

Envelope Id: 666318C9A78D44C196B3335ADB08F090
Status: Completed
Subject: Complete with DocuSign: 232SM203 Part B Later-onset SAP Final v1.0.docx
Vendor Name:
Source Envelope:
Document Pages: 59
Signatures: 3
Envelope Originator:
Certificate Pages: 2
Initials: 0
[redacted]
AutoNav: Enabled
225 Binney St
EnvelopeId Stamping: Enabled
Cambridge, MA 02142
Time Zone: (UTC) Dublin, Edinburgh, Lisbon, London
[redacted]
IP Address: [redacted]

Record Tracking

Status: Original
Holder: [redacted]
Location: DocuSign
11 April 2024 | 15:29
[redacted]

Signer Events	Signature	Timestamp
[redacted] [redacted] [redacted] Principal Biostatistician Biogen Security Level: Email, Account Authentication (None)	DocuSigned by: [redacted] Signature Adoption: Pre-selected Style Using IP Address: [redacted]	Sent: 11 April 2024 | 15:36 Viewed: 11 April 2024 | 15:36 Signed: 11 April 2024 | 15:36
Electronic Record and Signature Disclosure: Not Offered via DocuSign		
[redacted] [redacted] [redacted] Spinal Muscular Atrophy (SMA) Biogen Security Level: Email, Account Authentication (None)	DocuSigned by: [redacted] Signature Adoption: Pre-selected Style Using IP Address: [redacted]	Sent: 11 April 2024 | 15:36 Viewed: 11 April 2024 | 15:36 Signed: 11 April 2024 | 15:37
Electronic Record and Signature Disclosure: Not Offered via DocuSign		
[redacted] [redacted] [redacted] Biogen Security Level: Email, Account Authentication (None)	DocuSigned by: [redacted] Signature Adoption: Pre-selected Style Using IP Address: [redacted]	Sent: 11 April 2024 | 15:36 Viewed: 11 April 2024 | 15:47 Signed: 11 April 2024 | 15:50
Electronic Record and Signature Disclosure: Not Offered via DocuSign		

In Person Signer Events	Signature	Timestamp

Editor Delivery Events	Status	Timestamp

Agent Delivery Events	Status	Timestamp

Intermediary Delivery Events	Status	Timestamp

Certified Delivery Events	Status	Timestamp

Carbon Copy Events	Status	Timestamp

Witness Events	Signature	Timestamp

Notary Events	Signature	Timestamp

Envelope Summary Events	Status	Timestamps
Envelope Sent	Hashed/Encrypted	11 April 2024 | 15:36
Certified Delivered	Security Checked	11 April 2024 | 15:47
Signing Complete	Security Checked	11 April 2024 | 15:50
Completed	Security Checked	11 April 2024 | 15:50

Payment Events	Status	Timestamps

Signature Page

Document Name: 232SM203 Part B Later-onset SAP

Document Title: 232SM203 Part B Later-onset SAP

Signed by	Role	Date / Time (UTC)
[redacted]	Signing as Approver	11-Apr-2024 15:45:28

Signature Page

Document Name:

Document Title:

Signed by	Role	Date / Time (UTC)

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

PART C STATISTICAL ANALYSIS PLAN

Version No.: 1.0

Date: 06 October 2023

Author: [redacted]

Study Title:
Escalating Dose and Randomized, Controlled Study of Nusinersen (BIIB058) in Participants With Spinal Muscular Atrophy

Name of Study Treatment: Nusinersen

Protocol No.: 232SM203

Study Phase: 2/3

Confidential Information

The information contained in this Statistical Analysis Plan (SAP) is confidential and may not be reproduced or otherwise disseminated without the approval of Biogen, Inc.

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

APPROVAL

This document has been reviewed and approved by:		
[redacted]	DocuSigned by: [redacted] Signature	11-Oct-2023 Date
SMT Statistician		
[redacted]	DocuSigned by: [redacted]	11-Oct-2023 Date
Program Statistician		
[redacted]	DocuSigned by: [redacted]	11-Oct-2023 Date
SMT Medical Director		

VERSION HISTORY

SAP Version	Date	Primary Reasons for Amendment

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

Table of Contents

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

LIST OF ABBREVIATIONS

Abbreviation	**Definition**
[redacted]	[redacted]
[redacted]	[redacted]
aPTT	activated partial thromboplastin time
ACEND Assessment of Caregiver Experience with Neuromuscular Disease	Assessment of Caregiver Experience with Neuromuscular Disease
[redacted]	[redacted]
AE	Adverse event
ALT	alanine aminotransferase (SGPT)
AST	aspartate aminotransferase (SGOT)
AUC	Area under the curve
BLQ	Below limit of quantification
BUN	blood urea nitrogen
CGIC	Clinical Global Impression of Change
CHOP INTEND	Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders
[redacted]	[redacted]
CO2	bicarbonate
COVID-19	Coronavirus disease 2019
CSF	Cerebrospinal fluid
CV	Coefficient of variance

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

ECG	Electrocardiogram
eCRF	electronic Case report form
[redacted]	[redacted]
GGT	gamma-glutamyl transferase
HFMSE	Hammersmith Functional Motor Scale – Expanded
HINE	Hammersmith Infant Neurological Examination
HRQOL	Health-related quality of life
ICH	International Council for Harmonization
INR	International normalized ratio
[redacted]	[redacted]
IRT	Interactive response technology
ITT	Intent-to-Treat
LLOQ	lower limit of quantification
LP	Lumbar puncture
MET	Metabolic equivalent task
[redacted]	[redacted]
[redacted]	[redacted]
PT	prothrombin time
PedsQL	Pediatric Quality of Life Inventory™
PK	Pharmacokinetic(s)
PPS	Per protocol Set
QoL	Quality-of-life
QTcF	Corrected QT interval using Fridericia's formula

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

RULM	Revised Upper Limb Module
SAE	Serious adverse event
SD	Standard deviation
SE	Standard error
SMA	Spinal muscular atrophy
[redacted]	[redacted]
SMN	Survival motor neuron
SMN1	Survival motor neuron 1
SMN2	Survival motor neuron 2
SOC	System Organ Class
ULN	upper limit of normal
[redacted]	[redacted]
WHO	World Health Organization

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

1. Introduction

This Statistical Analysis Plan is based on Version 6 of the protocol, dated 05May2022. All references to the protocol refer to Version 6.

Nusinersen is an antisense oligonucleotide administered intrathecally via lumbar puncture (LP); it increases survival motor neuron (SMN) protein expression and significantly improves motor function in patients with spinal muscular atrophy (SMA). Nusinersen was approved for the treatment of SMA under the tradename Spinraza™ in the United States (US), European Union, and other countries. The population for this study includes participants with infantile-onset and later-onset SMA.

Efficacy and safety results across the nusinersen clinical development program have demonstrated an overall positive benefit-risk profile of nusinersen across a broad range of SMA phenotypes and patient populations. Nusinersen is approved in the US, Europe, and other countries and regions for the treatment of SMA in paediatric and adult patients at a recommended dosage of 12 mg administered in 3 loading doses at 14-day intervals, a fourth loading dose 30 days after the third dose, and maintenance doses every 4 months thereafter in most countries and regions. Pharmacokinetic (PK) and pharmacodynamic (PD) analyses indicate that nusinersen drug exposure higher than that achieved with 12 mg in patients with SMA may produce an even greater benefit in motor function. Additionally, PK modelling and simulations identified dosing regimens that achieve higher drug exposure more rapidly. Therefore, this study is being conducted to investigate the efficacy, safety, tolerability, [redacted] of a 50/28-mg dose of nusinersen (50-mg loading dose/28-mg maintenance dose) and a dosing regimen targeted to achieve higher drug exposure more rapidly. This study will be conducted in participants with genetically confirmed SMA.

This is a 3-part study Part A, Part B (infantile-onset and later-onset), and Part C.

Part A was an open label safety evaluation in which later-onset SMA subjects received 3 loading doses of 28 mg of nusinersen and 2 maintenance doses of 28 mg. All 6 participants enrolled in Part A completed the study.

In Part B, in order to evaluate the proposed higher dosing regimen, an active-controlled design is being used, with participants randomized either to the investigational dosing regimen (50-mg loading/28-mg maintenance) or to the currently approved dosing regimen (12-mg loading/12-mg maintenance). In order to supplement the number of participants available for the analysis in Part B of this study, historical data from the existing data set for Study CS3B, a Phase 3 efficacy and safety study of nusinersen in participants with infantile-onset SMA, may be borrowed to augment the control arm.

Part C of the study will allow safety evaluation of transitioning participants who are on the currently approved dose of nusinersen (12-mg maintenance for at least 1 year after the initiation of treatment) to the proposed higher dosing regimen via the administration of a single bolus dose

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

of 50 mg of nusinersen (which should be administered 4 months ± 14 days after their most recent maintenance dose of 12 mg), with maintenance dosing at 28 mg thereafter.

This SAP will cover the planned analysis of Part C.

2. Study Overview

2.1. Study Objectives and Endpoints

Study Primary Objective (Part C)

To examine the safety and tolerability of nusinersen administered intrathecally at higher doses to participants with SMA.

Study Primary Endpoints (Part C)

- Incidence of adverse events (AEs), including serious adverse events (SAEs)
- Shifts from baseline in clinical laboratory parameters, electrocardiograms (ECGs), and vital signs
- Change in growth parameters
- Shifts from baseline in coagulation parameters (activated partial thromboplastin time [aPTT], prothrombin time [PT], and international normalized ratio [INR])
- Change in urine total protein
- Change from baseline in neurological examination outcomes
- The proportion of participants with a post-baseline platelet count below the lower limit of normal on at least 2 consecutive measurements
- The proportion of participants with a post-baseline QT interval using Fridericia's formula (QTcF) of > 500 msec and an increase from baseline to any post-baseline timepoint in QTcF of > 60 msec

Study Secondary Objective (Part C)

- To examine the clinical efficacy of nusinersen administered intrathecally at higher doses to participants with SMA.
- To examine the effect of nusinersen administered intrathecally at higher doses to participants with SMA

Study Secondary Endpoint (Part C)

- Change from baseline in Hammersmith Functional Motor Scale Expanded (HFMSE) score
- Change from baseline in Revised Upper Limb Module (RULM) score
- Total number of new World Health Organization (WHO) motor milestones
- Change from baseline in Assessment of Caregiver Experience with Neuromuscular Disease (ACEND)
- Change from baseline in Pediatric Quality of Life Inventory™ (PedsQL)

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

- Change from baseline in Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders (CHOP-INTEND) (Cohort 1 only)
- Change from baseline in Hammersmith Infant Neurological Examination (HINE) Section 2 motor milestones (Cohort 1 only)
- Number and duration of hospitalizations
- Clinical Global Impression of Change (CGIC) [physician, caregiver] at Day 302
- Number of serious respiratory events
- Ventilator use

[redacted]

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

[redacted]

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

2.2. Study Design

This 3-part study will evaluate the efficacy and safety of a higher dosing regimen of nusinersen in approximately 145 participants. The study will be conducted at approximately 65 sites

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

globally. Following the completion of this study, all eligible participants may elect to enrol in a separate long-term extension study (232SM302), pending study approval by ethics committees and the appropriate regulatory authorities. In regard to participants rolling over from Part B, this will be done without unblinding the participant's treatment group.

In Part C, up to approximately 40 participants who have already initiated treatment with nusinersen and have been receiving the approved dose of 12 mg for at least 1 year prior to entry will be enrolled in Part C. The initial cohort in Part C (i.e., Cohort 1) consists of up to approximately 20 participants of any age and of any SMA status. For Cohort 1, an attempt will be made to enroll at least 8 but no more than 12 participants ≥ 18 years of age (participants in Cohort 1 ≥ 18 years of age must be ambulatory). Up to 5 participants with severe scoliosis and/or severe contractures may be enrolled in Cohort 1 of Part C after consultation with the Medical Monitor. An additional cohort (i.e., Cohort 2) consisting of up to approximately 20 adult participants (≥ 18 years of age) was subsequently added to Part C in Protocol Version 5, in order to enable collection of data in adults transitioning from the currently approved nusinersen dosing regimen to a higher dose. Participants in Cohort 2 can be either ambulatory or non-ambulatory. All participants in Part C will receive a single bolus dose of 50 mg (which should be administered 4 months ± 14 days after their most recent maintenance dose of 12 mg) followed by 2 maintenance doses of 28 mg on Days 121 and 241. Participants in Part C will remain at the clinic for at least 24 hours after the first (bolus) dose for the purpose of completing study assessments.

2.3. Sample Size Considerations

For Cohort 1 of Part C, an attempt was made to enroll at least 8 but no more than 12 participants ≥ 18 years of age (participants ≥ 18 years of age must be ambulatory). For Cohort 2 of Part C, all participants must be ≥ 18 years of age and can be either ambulatory or non-ambulatory.

2.4. General Trial Conduct Mitigation Strategy under COVID-19 Pandemic or humanitarian emergencies

In order to mitigate risk of missing dosing or assessments during the time of the COVID-19 pandemic, the following has been put into place:

- If a site is not closed due to COVID-19 or humanitarian emergencies and the site enables the participant to attend the visit in-clinic, then delayed visits within a reasonable timeframe are allowed.
- Site transfers are encouraged where possible, so that subjects can be transferred to another site for assessment. Instructions are in place to enable sites to do this including transfer of the database and source documents for the patient from the transferring site to the receiving site. This will allow the participant to receive dosing and protocol defined assessments to be made. Dosing can only be performed in the clinic setting due to intrathecal administration. In particular, if screening cannot be performed in the clinic, participants will be screened at a screening site and subsequently transferred to another

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

site to perform dosing and all assessments for the participant, since dosing can only be performed in the clinic setting due to intrathecal administration.

- A phone visit can be used to collect safety data (AEs, concomitant therapies/procedures)
- Transferring sites are encouraged to enter and clean all data before transferring the participant to the receiving site. The database for a participant will remain with the transferring site until data are entered and cleaned as much as possible, but the transferring site must have access to the participant in IRT and copies of all source documents. Any queries raised later on for data belonging to visits at the transferring site will be cleaned by the receiving site through the source documents, and interaction between monitors at the receiving and transferring sites. The transferring site will no longer have access to the database for a participant they have transferred, but the monitors will. The Subject ID will remain the same when a participant is transferred to another site, but the transfer site can be tracked via the Site ID.

3. Definitions

3.1. Dates and Points of Reference

In the remainder of this analysis plan only Part C analyses will be described. Separate SAPs will be developed for other parts of the study, (Parts A open label and B randomized double-blind). Therefore 'Part C' will not be stated throughout this document.

- Study Day 1: the date of the first dose of study treatment in 232SM203
- Study Day
  - For a date on or after Study Day 1

    Study Day = (Date of Interest) – (Study Day 1) + 1
  - For a date before Study Day 1

    Study Day = (Date of Interest) – (Study Day 1)

In order to distinguish nominal visit names from duration defined in days, visit names will be referred to as "Day 121", "Day 241", etc., and "121 days" or "241 days", etc. will be used to define time intervals.

Overall time on commercial nusinersen prior to first study dose= (Date of first study dose) – (Date of first commercial dose)

Overall time on study will be defined as the total number of days a participant is known to be followed on study calculated as follows:

Overall time on Study = (Last date on study) – (Date of first study dose) +1

Last date on study is defined as the date of the latest visit or evaluation or telephone contact, or time of death from all available data for a given participant.

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

Disease duration at informed consent is defined as time from age at symptom onset to age at informed consent.

3.2. Study Treatment

All participants will receive a single bolus dose of 50 mg of nusinersen administered intrathecally on Day 1 of this study (which should be administered 4 months ± 14 days after their most recent maintenance dose of 12 mg), followed by 2 maintenance doses of 28 mg of nusinersen on Days 121 and 241.

3.3. Study Groups

All the summaries will be using display 'a' below, only selected analysis will be repeated using displays 'b' and 'c'and will include baseline characteristics, disposition, exposure, time on study incidence of adverse events and secondary endpoints.

a.

Nusinersen 50/28 mg			
Infantile-onset SMA	**Later-onset SMA**		**Total**
	Age < 18 years	Age>= 18 years	

b.

Nusinersen 50/28 mg					
Non-Ambulatory			**Ambulatory**		
Age <18 years	Age >=18 years	Total	Age< 18 years	Age>= 18 years	Total

c.

Nusinersen 50/28 mg					
< median time on commercial nusinersen prior first dose			**>= median time on commercial nusinersen prior first dose**		
Age <18 years	Age >=18 years	Total	Age < 18 years	Age >= 18 years	Total

3.4. Study Periods

The total study duration for each participant will be a maximum of 323 to 382 days with 5 to 6 visits.

- Screening: 21 days
- Loading period: 1 day
- Maintenance period: 240 days

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

- Follow-up: 61 to 120 days

3.5. Key Derived Variables

Not applicable.

3.6. Stratification Factors and Subgroup Variables

3.6.1. Stratification Factors

Not applicable.

3.6.2. Subgroup Variables

Not applicable.

3.7. Analysis Sets

The Intent-to-Treat (ITT) Set is defined as all participants who receive at least one dose of nusinersen in 232SM203.

The Safety Analysis will be performed on the ITT Set.

[redacted]

A Per Protocol Set (PPS) will include the subset of the ITT Set who have no significant protocol deviations that would be expected to affect efficacy assessments

Significant protocol deviations will be determined prior to database lock and will include:

Participants with missing baseline results for the following assessments: HFMSE or RULM.

Participants who took experimental or approved agents for the treatment of SMA, including gene therapy, biological agent, any disease modifying therapy, or device, during the study. This includes marketed agents at experimental doses that are being tested for the treatment of SMA (oral albuterol/salbutamol, riluzole, carnitine, sodium phenylbutyrate, valproate, hydroxyurea, etc.).

4. List of Planned Study Analyses

4.1. Interim Analysis

No Interim analysis is planned.

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

4.2. Primary Analysis

Not applicable.

4.3. Final Analysis

The final analysis will be conducted after all Part C subjects have completed the study.

5. Statistics Methods for Planned Analyses

5.1. General Principles

Descriptive summary statistics will be presented for all primary, secondary [redacted] endpoints collected. Unless otherwise specified, for continuous endpoints, the summary statistics will generally include number of participants with data, mean, standard deviation, median, 25th percentile, 75th percentile, minimum and maximum. For categorical endpoints, the summary statistics will generally include number of participants with data, and the percentage of those with data in each category.

All summaries and listings will be presented by study groups unless otherwise specified. Visits in listings will be displayed as per CRF data collection rather than analysis visits.
The statistical software, SAS®(Version 9.4) will be used for all summaries and statistical analyses.

Visit Windows for Early Withdrawal, Unscheduled Visits

Data from early withdrawal visits and post baseline unscheduled assessments will be mapped to a planned visit using a windowing scheme for assessments that are tabulated by visit. Scheduled visits will not be windowed.

The visit windowing will be performed for safety and efficacy endpoints as shown below.

Table 2: Visit Windows for safety and efficacy Endpoints in 232SM203:

Visit	Lower Bound	Upper Bound	Target Day
Baseline			<=1
Day 121	2	181	121
Day 241	182	271	241
Day 302	272	422	302

Any postdose assessments with specific time associated to it (i.e vital signs collected at 1,2,4,6 and 24 hours) will be associated to the windowed visit at which they belong.

The following rules will be implemented for windowing of these visits for both safety and efficacy.

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

- If more than one observation is within the same window, data from the regular scheduled visit will be used for that visit.
- If neither of the observations are from a regular scheduled visit the observation closest to the planned target date will be used. However, if both the observations are equidistant from the target date, the latest one will be used; if repeated measurements are on the same day, then the last measurement will be used.
- If there is more than one observation in a window for a dosing visit the observation on the day of dosing will be chosen over the observation where dosing did not occur on the same day.
- In windowing laboratory data, if a central and local result is available within the window then the central laboratory result will be chosen

5.2. Participant Accountability

The number (and percentage) of subjects screened, enrolled, dosed, completed treatment and study, along with the reasons for discontinuing treatment and withdrawing from the study, will be presented.

If there are any subjects who discontinued from treatment or withdrew from study due to reasons associated with the COVID-19 pandemic or reasons associated with humanitarian emergencies, a separate summary will be presented to summarize these reasons. Adverse events, deaths, and other reasons may fall into COVID-19 related categories.

A listing of subjects who discontinued treatment/withdrew from study and the associated reasons for discontinuation/withdrawal will be presented. Subjects who died during the study will be listed separately. Two separate listings will also be presented for subjects who discontinued treatment/withdrew from study due to the COVID-19 pandemic or to humanitarian emergencies if applicable.

5.3. Demographic and Baseline Characteristics

Baseline data (demography, medical history, SMA history, and baseline disease characteristics) will be summarized.

Demography includes age at screening, age at first dose in the study, sex, ethnicity, and race. Medical history will be coded in MedDRA and the number and percentage of subjects with each medical history presented by preferred term.

Demographic, baseline disease characteristics will be presented for the ITT Set and PPS.

SMA history includes age of symptom onset, time from disease onset to enrollment, age at SMA diagnosis, time from diagnosis to enrollment, number of copies of the SMN1 and SMN2 gene, highest motor function achieved, wheelchair use, and walker/cane/crutches/braces use.

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

Baseline disease characteristics presented include HFMSE, WHO motor milestones, Revised Upper Limb Module Test, growth parameters, CHOP INTEND, COBB angle and Neurological Examination (HINE) Section 2 motor milestone if available. These assessments are further described in the next paragraph and Section 5.

Commercial nusinersen prior first dose

The number of commercial nusinersen doses received, prior to the first study dose, will be displayed using frequency distributions. The cumulative amount of commercial nusinersen received, will be summarized using summary statistics. The overall time on commercial nusinersen will be summarized descriptively (see section 3.1 for calculation).

5.4. Protocol Deviations

Protocol deviations identified during site monitoring will be captured in a protocol deviation log and categorized as major or minor deviations. These deviations will be listed. Major protocol deviations will be summarized including those that lead to exclusion from the Per protocol Set.

5.5. Study Treatment Exposure and Concomitant Medications

Exposure on Study

The number of doses received will be displayed using frequency distributions. The cumulative amount of nusinersen received will be summarized using summary statistics.

The overall time on study will be summarized descriptively (see section 3.1 for calculation). The duration will also be categorized and summarized using the following categories: <121 days, >=121 to 241 days, >=241 to 302 days, >=302 days

Given the long half-life of nusinersen, participants are considered to be exposed to the date of last visit or contact.

Separate listings will be provided showing study drug administration data which will include lot numbers, actual treatment received, cumulative number of doses and cumulative doses.

The cumulative, prior and during study, time on nusinersen will also be summarized as above.

Concomitant Medications and Procedure

A concomitant therapy is any drug or substance administered between screening and the final study visit/telephone call. Participants should be instructed to contact their Investigators before taking any new medications, including nonprescription drugs and herbal preparations.

Throughout the study, Investigators or designated licensed physicians involved in the study may prescribe concomitant therapies (procedures) or treatments deemed necessary for AEs or to provide adequate supportive care.

Participants are prohibited from receiving other experimental or approved agents for the treatment of SMA, including gene therapy, biological agent, or device, during the study. This

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

includes marketed agents at experimental doses that are being tested for the treatment of SMA (oral albuterol/salbutamol, riluzole, carnitine, sodium phenylbutyrate, valproate, hydroxyurea, etc.).

All concomitant medications will be coded using the World Health Organization drug dictionary (WHO Drug).

A concomitant procedure is any therapeutic intervention (e.g., surgery/biopsy, physical therapy) or diagnostic assessment (e.g., blood gas measurement, bacterial cultures) performed between Screening and the final study visit/telephone call. Concomitant procedures will be coded using MedDRA.

For the purposes of analysis, a concomitant therapy (including medication or procedure) is defined as any therapy that was taken or administered on or after the first injection of nusinersen received in the study. This includes therapies that were started prior to the initiation of injection of nusinersen if their use continued on or after the first injection of nusinersen received in the study. Refer to Appendix A for more details about missing start or stop dates.

The number and percentage of subjects who were taking each type of concomitant medication at baseline and during the study will be presented. The number and percentage of subjects taking each type of concomitant procedure will be presented by preferred term.

If there are a sufficient number of tests or treatments reported with indication of COVID-19, a separate summary of these will be provided.

Concomitant medications will also be reviewed by a medical reviewer prior to the final lock to determine disallowed medications according to the protocol. These will be summarized separately.

5.6. Efficacy Endpoints

5.6.1. General Analysis Methods for Efficacy Endpoints

The following secondary [redacted] efficacy assessments will be evaluated:

- HFMSE
- RULM
- WHO Motor Milestones
- ACEND
- PedsQL
- Hospitalizations
- CGIC
- Serious respiratory events
- Ventilator use
- [redacted]

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

Cohort 1 only:

- CHOP INTEND
- HINE Section 2 motor milestones

[redacted]

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

Unless otherwise specified, the baseline for the clinical function endpoints and QoL questionnaires are defined as the last non missing assessment prior to the first dose of study treatment.

5.6.2. Imputation

Multiple imputation (MI)

In the event more than two subjects have missing data for an endpoint at a post baseline visit then multiple imputation will be performed. If applicable imputed data summaries will be presented and also the summaries of observed data.

For multiple imputation method [Schafer 1997, Schafer 1999] will be used for patients who discontinue for a reason other than death or have a missing efficacy result at the timepoint of interest.

For HFMSE and RULM the imputations will be performed on total scores for post baseline visit. The imputations will be performed on total scores for post baseline visits. Missing baseline will not be imputed. All available data will be used i.e., where subjects have missed doses but continued to perform assessments; otherwise, any missing values will be imputed using MI.

The Markov Chain Monte Carlo (MCMC) method will be used.

Prior to the Proc MI step the dataset will be sorted by ascending unique subject identifier (USUBJID). The variable list in the model for imputations will include:

- Age at onset of symptoms, age at first dose, baseline ambulatory status, baseline value for the endpoint, and all available post baseline values

A set of 100 complete imputed datasets will be generated and the relative efficiency parameter will also be checked to determine the acceptability of the imputed results i.e. it should be close to 98% or higher.

For continuous variables, each of the 100 imputed datasets will be used to calculate the change from baseline and subsequently the correspondent means, and standard errors. These summaries

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

will be used to calculate a combined estimate of mean and mean change from baseline using PROC MIANALYZE [Little et al, 2002] for each visit.

In the presentation of results from multiple imputed data, the number of subjects with missing data and the degree to which it is missing will be summarized.

Imputation for WHO motor milestones

If for a milestone either 'No (refusal)' or 'Unable to test' are observed at a visit then the result will be first set to missing. Imputation will be performed for missing data considering each milestone separately using the following rules for scheduled visits.

For baseline, the closest non-missing milestone prior to or at first dose will be selected. If this results in a subject having no baseline for all six milestones, then this will be left missing, and no imputation will be performed. If at least one milestone is evaluable as either Yes or No and the remainder are missing, then other assessments or SMA history will be used to impute a baseline result.

For post baseline visits, if a subject has at least one milestone assessed but the remainder are missing the missing value will be imputed by using the worst result from the flanking visits. Otherwise, if the imputation is the last visit, the missing value will be imputed as the last observed value for the subject. Of note, only observed data will be utilized for imputation purposes and in a situation where unscheduled visit data is available these values will be utilized if these flank a missing visit. Missing motor milestone items will be imputed first prior to any analysis.

5.6.3. Primary Efficacy Endpoints

Not applicable.

5.6.4. Secondary Efficacy Endpoints

For each endpoint the main analysis will be performed on the ITT set and be repeated for PPS Set.

5.6.4.1. HFMSE

The HFMSE is a tool used to assess motor function in children 2 years or older with SMA. The scale was originally developed with 20 scored activities and was devised for use in children with Type 2 and Type 3 SMA with limited ambulation to give objective information on motor ability and clinical progression [Main 2003]. The expanded scale includes an additional module of 13 items developed to alleviate the ceiling effect and allow evaluation of higher functioning and ambulatory SMA patients [O'Hagen 2007]. Each item is scored 0 (unable), 1 (performs with modification or adaption) or 2 (able) and the total score is calculated by summing the 33 items and ranges from 0 to 66 with higher scores indicating greater motor function.

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

If 6 or fewer items are missing, then these items will be imputed to 0 when summing all 33 items. If greater than 6 items are missing, then the total score will be set to missing.

In order to further explore the response as measured by HFMSE score a number of thresholds will be evaluated: The proportion of subjects achieving: worsening of >= 4, >= 3, >=2, >=1 points, any worsening, no change (=0 points), any improvement, >=1, >=2, >=3, >=4 points improvement will be presented.

Analyses:

- Total score and change from baseline by visit
- Thresholds of change from baseline by visit
- Thresholds of change from baseline to last observed assessment
- Mean (-/+ SE) plots of total score and change from baseline over time

Sensitivity analyses

- The above analyses will be repeated excluding subjects with a baseline cobb angle >=40.
- Total score and change from visit displays will be repeated including patients with baseline HFMSE consistent with the CS4 study: baseline HFMSE >=10 and <= 54

5.6.4.2. *RULM*

The upper limb module test (Mazzone et al. 2011) was developed to assess the upper limb functional abilities in SMA patients using 9 items. However, a revised version of the upper limb module test developed by the Upper Limb Module Working Group will be performed in this study consisting of 20 items. The first item is assessed on a seven point scale ranging from 0 – "No useful function" to 6 – "Can abduct both arms simultaneously elbows in extension in a full circle until they touch above the head". This first item will not contribute to the total score. For the remaining 19 items, 18 items are assessed on a three point scale using the following criteria: 0 – Unable to achieve independently; 1 – Modified method but achieves goal independent of physical assistance from another person; 2 – Normal – achieves goal without any assistance. The one remaining item is assessed as either 0 or 1. For each item, a score will be collected on the left and right side. A derived total score will be calculated by summing the scores from these 19 individual items and ranges from 0 if the subject fails all activities to 37 if the subject achieves all activities. If, for an individual item, a response is recorded for both he left and right side the highest score will be used in calculating the total. If 3 or fewer items are still missing responses then it will be assumed that the score was 0. If greater than 3 items are missing, then the total score will be set to be missing.

Missing baseline will not be imputed.

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

In order to further explore the response as measured by upper limb score a number of thresholds will be evaluated: The proportion of participants achieving: worsening of >= 4, >= 3, >=2, >=1 points, any worsening, no change, any improvement, >=1, >=2, >=3,>=4 points improvement will be presented.

Analyses

- Total score and change from baseline by visit
- Thresholds of change from baseline by visit
- Thresholds of change from baseline to last observed assessment
- Mean (-/+ SE) plots of total score and change from baseline over time

Sensitivity analyses

- The above analyses will be repeated excluding subjects with a baseline cobb angle >=40 .
- Total score and change from visit will be repeated including patients with baseline RULM: >=5 and <= 30

5.6.4.3. *WHO motor milestones*

The WHO motor milestones criteria [WHO Multicentre Growth Reference Study Group 2006; Wijnhoven 2004] are a set of six milestones in motor development, all of which would be expected to be attained by age 24 months in healthy children. The individual milestones are:

- Sitting without support
- Standing with assistance
- Hands and knees crawling
- Walking with assistance
- Standing alone
- Walking alone

As part of the assessment, the examiner records an overall rating of the subject's emotional state and then for each milestone one of the following four classifications:

- No (inability) – Child tried but failed to perform the milestone
- No (refusal) – Child refused to perform despite being calm and alert
- Yes – Child was able to perform the milestone
- Unable to test – Could not be tested because of irritability, drowsiness or sickness

WHO motor milestones will be assessed by the site clinical evaluator and caregiver. Adult

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

participants who do not require a caregiver during the study visit may self-report WHO motor milestone achievement. Whenever possible, for each participant, motor milestone assessments across all study visits should be conducted consistently by the same clinical evaluator.

If for a milestone either 'No (refusal)' or 'Unable to test' are observed at a visit then the result will be set to missing.

New motor milestones

For each milestone at a visit and baseline values will be compared and if the subject can achieve a new milestone compared to baseline then it will be counted +1, if the subject is unable to achieve the milestone then it will be scored -1 and if the milestone has been maintained then it will be scored 0. These scores will then be summed to give a total score which could range between -6 (if a subject was able to perform all milestones at baseline but lost the ability and could not perform any milestones at the specific visit) and + 6 (if a subject was unable to achieve any milestones at baseline but could achieve all at the specific visit).

Analyses

- Summary of motor milestones at baseline
- Number of new motor milestones achieved per subject by visit
- Summary of motor milestones status at last observed assessment

5.6.4.4. *PedsQL*

Subjects who are ≥ 2 years of age in cohort 1 at Screening will be evaluated using the PedsQL Measurement 4.0 Generic Core Scales and 3.0 Neuromuscular Module [Varni 1999]. This assessment has been validated by the American Spinal Muscular Atrophy Randomized Trials Group [Iannaccone 2009].

The PedsQL Measurement Model is a modular approach to measuring health-related quality of life (HRQOL) in children, adolescents and adults. The PedsQL consists of brief, practical, generic core scales, as well as condition-specific modules for use in designated clinical populations. The PedsQL4.0 Generic Core Scales include an assessment of physical functioning, emotional functioning, social functioning, and school functioning and will be assessed for participants 2 to ≥ 26 years of age. The PedsQL 3.0 Neuromuscular Module was designed to measure HRQOL dimensions specific to patient with neuromuscular disorders, including SMA, and will be assessed in participants 2 to 25 years of age. Patient self-report will be measured in participants starting at 5 years of age, while parent proxy-report of HRQOL will be measured for participants starting at 2 years of age.

If a participant is assessed on PedsQL at baseline, then the age range of the scale used at Screening for that participant will continue to be used up to Day 302, regardless of changes in the participant's age.

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

The PedsQL Core Scales parent questionnaire is collected for children in the following age categories: 2-4, 5-7, 8-12, 13-18, 18-25, and ≥26. Four dimensions are collected: Physical, Emotional, Social and School functioning and each item is scored on a 5-point ordinal scale (0= Never, 1 = Almost Never, 2= Sometimes, 3 = Often, 4 = Almost Always).

The PedsQL Core Scales patient questionnaire is collected for children in the following categories: 5-7, 8-12, 13-18, 18-25, and ≥26. Similar dimensions and 5-point ordinal scale are used as for the parents but for subjects aged 5 to 18 years a 3-point ordinal scale is collected, omitting the response levels of 1 and 3.

In the neuromuscular module, one parent questionnaire is collected for all subjects irrespective of age with three dimensions: 'About my child's neuromuscular disease', 'Communication' and 'Family resources'. The same 5-point ordinal scale is collected for each question.

The patient neuromuscular disease questionnaire is collected for subjects in the following age categories: 5-7, 8-12, 13-18 and 18-25. The questionnaire for subjects aged 5-7 years uses the 3-point ordinal scale as above and has only one dimension - 'About my Neuromuscular disease'. The questionnaires for other age categories use 5-point ordinal scale with three dimensions.

In scoring a dimension (physical functioning, emotional functioning, social functioning, and school functioning) the first step is to reverse and linearly transform to a 0-100 scale (0 = 100, 1 =75, 2=50, 3= 25, 4= 0), so a higher score is indicative of a better health related quality of life. If greater than 50 percent of the items within a dimension are missing then the dimension score will not be computed, otherwise the mean score for the dimension will be calculated as the sum of items over the number of items answered.

A psychosocial health summary score, constructed from three dimensions, will be calculated as the sum of items over the number of items answered in the emotional, social and school functioning scales. A total score will be calculated as the sum of all the items over the number of items answered on all the scales. If greater than 50 percent of the items are missing, then the summary score or total score will be set to be missing.

For the neuromuscular module, a score for each dimension and then total score will be calculated in the same manner, no health summary scores are evaluated.

Due to the age specific nature of these questionnaires, subjects aged 2-4 years would not be expected to complete the self-evaluation.

Analyses

- Total score and change from baseline by visit
- Total score and change (parent/subject evaluation separately) from baseline by visit
- Mean (-/+ SE) plots of total score and change from baseline.

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

5.6.4.5. *ACEND*

Parents/caregivers of participants will complete the ACEND questionnaire. The ACEND questionnaire will not be collected for adult participants who do not require a caregiver during the study visits. For cohort 1, only the parents/caregivers of the participants older than 2 years will complete this questionnaire. This assessment instrument has been designed to quantify the caregiver impact experienced by parents/caregivers of children affected with severe neuromuscular diseases, including children with SMA [Matsumoto 2011]. ACEND includes seven domains assessing physical impact (including feeding/grooming/dressing, sitting/play, transfers, and mobility) and general caregiver impact (including time, emotion, and finance). and each domain comprises several items.

The total score for a domain with n items, each item assessed on ordinal scale from 1 to *z*, is derived as follows: 100 multiplied by (Mean of the n items in the domain -1) divided by (*z*-1).

This total score will be on a scale of 0 to 100 with a higher score indicating a greater impact on the caregiver.

At least two items for the time domain and one item for the remaining domains need to be non-missing for a total to be calculated; else the total score will be set to be missing.

Analyses

- Total score and change from baseline by visit
- Domain total score and change from baseline by visit
- Mean (-/+ SE) plots of total score and change from baseline.

5.6.4.6. CHOP INTEND

The CHOP INTEND test was specifically designed to evaluate the motor skills of infants with significant motor weakness, including infants with SMA [Glanzman 2010]. The CHOP INTEND test captures neck, trunk, and proximal and distal limb strength in 14 elicited and 2 observational items. The CHOP INTEND has been established as a safe and reliable infant motor measure in infantile-onset SMA and has been validated [Glanzman 2011].

The CHOP INTEND infant motor function scale is comprised of 16 test items, nine of which are scored 0, 1, 2, 3, or 4 with greater scores indicating greater muscle strength, five are scored as 0, 2, or 4, one is scored as 0, 1, 2, or 4, and one as 0, 2, 3, or 4. This can result in a worst possible total score of 0 to a best possible total score of 64. CHOP INTEND is used to assess spontaneous movement in the upper extremities, spontaneous movement in the lower extremities, hand grip, head in midline with visual stimulation, hip adductors, rolling elicited from the legs, rolling

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

elicited from the arms, shoulder and elbow flexion and horizontal abduction, shoulder flexion and elbow flexion, knee extension, hip flexion and foot dorsiflexion, head control, elbow flexion, neck flexion, head/neck extension, and spinal incurvation. CHOP INTEND is to be evaluated during the study for cohort 1. For each item, a score will be collected on the left and right side.

In order to examine if the response is consistent with a range of thresholds, the proportion of participants achieving: worsening of >=6, >=5, >=4, >=3, >=2, >=1 points, no change (=0 points), and improvement of >=1, >=2, >=3, >=4 (endpoint for main analysis), >=5, >=6 points will be presented. If a subject has moved scales and/or reached maximum score of 64 will be summarized on a different row in the tables.

Depending on data availability the following displays will be presented showing the following

- Total score and change from baseline by visit
- Thresholds of change from baseline by visit

5.6.4.7. HINE Section 2

Hammersmith Infant Neurological Examination (HINE) is collected for cohort 1 infantile-onset SMA subjects, that are less than 2 years of age at the time of informed consent.

The assessment is comprised of eight motor milestone categories: voluntary grasp, ability to kick in supine position, head control, rolling, sitting, crawling, standing and walking.

Within each motor milestone category (depicted as Motor Milestones in Table 5), there are 3 to 5 levels that can be achieved (depicted as Milestones Progression, or rows, in the Table above). All 8 Motor Milestones are tested during each assessment. A subject whose results after testing all appear in the first column (No grasp, No kicking, Unable to maintain head upright, etc.) has not achieved any motor milestone. Motor milestone achievement is depicted by movement from the left side of the Table above to right side of the Table, as denoted by the Milestone Progression arrow in the Table.

The 8 categories of HINE Section 2 can be summed to give a total score that ranges from 0 to 26.

Depending on data availability the following displays:

- Total motor milestone score and change from baseline by visit
- Summary of highest motor milestones item achieved at last observed assessment.
- The proportion of subjects who have achieved specific individual item at last analysis visit.
- Mean (-/+ SE) plots of total score and change from baseline.

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

Table 5:Hammersmith Infant Neurological Examination Section 2 - Motor Milestones

Motor Milestone Category	Milestone Level Progression Score (Age Expected in Heathy Infants[1]) → Improvement				
	0	**1**	**2**	**3**	**4**
Voluntary grasp	No grasp	Uses whole hand	Finger and thumb/ immature grasp	Pincer grasp	
Ability to kick (in supine)	No kicking	Kicks horizontal; legs do not lift	Upward (vertically) [3 months]	Touches leg (4 to 5 months)	Touches toes (5 to 6 months)
Head control	Unable to maintain upright (< 3 months)	Wobbles (4 months)	All the time upright (5 months)		
Rolling	No rolling	Rolls to side (4 months)	Prone to supine (6 months)	Supine to prone (7 months)	
Sitting	Cannot sit	Sits with support at the hips (4 months)	Props (6 months)	Stable sit (7 months)	Pivots (rotates) [10 months]
Crawling	Does not lift head	On elbow (3 months)	On outstretched hand (4 to 5 months)	Crawls flat on the abdomen (8 months)	On hands and knees (10 months)
Standing	Does not support weight	Supports weight (4 to 5 months)	Stands with support (8 months)	Stands unaided (12 months)	

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

Motor Milestone Category	**Milestone Level Progression Score (Age Expected in Heathy Infants[1])** → **Improvement**				
	0	**1**	**2**	**3**	**4**
Walking	No walking	Bouncing (6 months)	Cruising (holding on) [11 months]	Walks independently (15 months)	

Values for healthy infants in [De Sanctis 2016; Haataja 1999]

5.6.4.8. Number and Length of Hospitalizations

The number of hospitalizations during the study will be analyzed using the rate at which they occur. For descriptive purposes, the aggregate hospitalization rate will be calculated for each group by dividing the total number of hospitalizations that occurred in the group by the total number of subject-years on study.

Annualized hospitalization rate will be calculated for each subject as the number of hospitalizations that the subject experienced divided by the number of days on study and this ratio multiplied by 365.

For each subject, the total time spent in hospital during the study will be calculated and prorated according to number of days on study.

5.6.4.9. CGIC

The CGIC rating scale was developed as a brief standalone assessment of the clinician's view of the patient's global functioning after initiating a study medication [Guy 1976]. The CGIC provides an overall clinician-determined summary measure that takes into account all available information, including knowledge of the patient's history, psychosocial circumstances, symptoms, behavior, and the impact of the symptoms on the patient's ability to function. The CGIC is administered by an experienced clinician who is familiar with the disease under study and the likely progression of treatment. The clinician makes a judgment about the total picture of the patient at each visit: the severity of the illness, the patient's level of distress and other aspects of impairment, and the impact of the illness on functioning. The CGIC is rated without regard to the clinician's belief that any clinical changes are or are not due to medication and without consideration of the etiology of the symptoms.

The CGIC is a 7-point scale that requires the clinician to assess how much the patient's illness has changed relative to a baseline state at the beginning of the intervention and is rated as follows: 1, very much improved; 2, much improved; 3, minimally improved; 4, no change; 5,

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

minimally worse; 6, much worse; or 7, very much worse. CGIC should be assessed consistently by the same rater for each study participant. A separate CGIC assessment will be performed by the Investigator (Principal Investigator or Sub investigator) and caregiver.

In order to investigate response on these scales, three definitions of responder will be defined: "Much improved": Responder (<=2) versus Non-responder (>=3), "Any improvement": Responder (<=3) versus Non-responder (>=4), "No worsening": Responder (<=4) versus Non-responder (>=5).

Analyses

- Proportion of participants achieving each responder definitions above by visit

5.6.4.10. Number of Serious Respiratory Events

All adverse events that are coded into the System Organ Class (SOC) of respiratory, thoracic, and mediastinal disorders, either as their primary SOC or their secondary SOC will be considered respiratory events. Only those classified as serious will be analyzed. By utilizing the above approach, rare occurrence of serious respiratory events in the Infection and Infestation SOC and Investigation SOC (if any) may not be included. However, a detailed review of the AE listing will be conducted to identify such events.

The number of serious respiratory events during the study will be analyzed using the rate at which they occur. For descriptive purposes, the aggregate rate will be calculated for each group by dividing the total number of serious respiratory events that occurred in the group by the total number of subject-years on study.

In order to evaluate if there is any change in the number of serious respiratory events overtime, the events will be presented in periods of 120 days.

5.6.4.11. Ventilator Use

The participant's use of ventilators and respiratory aids will be collected at every study visit.

The ventilation/respiratory aid use will be summarized by visit and listed.

[redacted]

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

[redacted]

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

[redacted]

[redacted]

[redacted]

- [redacted]
- [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

- [redacted]
- [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

- [redacted]

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

- [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

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

[redacted]

[redacted]

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

[redacted]

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

[redacted]

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

[redacted]

5.7. Safety Endpoints

Analyses of safety data will include adverse events and serious adverse events, laboratory data, ECGs, growth parameters, vital signs, and neurological examinations. Worsening or new

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

findings noted from the physical or neurological examinations will be reported as AEs as appropriate.

5.7.1. Adverse events

All adverse events (AEs) will be analyzed based on the principle of treatment emergence. An adverse event will be regarded as treatment-emergent if it was present prior to receiving the first dose of nusinersen in 232SM203 and subsequently worsened in severity or was not present prior to receiving the first dose of nusinersen and subsequently appeared. Refer to Appendix A for more details about AE with missing start or stop dates.

The incidence and frequency (event count) of treatment-emergent adverse events will be summarized. A participant having the same adverse event more than once will be counted only once in the incidence for that adverse event; multiple occurrences of the same adverse event for the same subject will all be counted in the frequency for that adverse event. Incidence will be presented by decreasing order by system organ class and by decreasing order by preferred term within each system organ class.

The occurrence of the AE with the greatest severity will be used in the calculation of incidence by severity. The AE relationship to treatment or LP procedure is captured as related or not related. In the event the relationship is missing then this will be considered related.

Due to the long half-life of nusinersen, analyses of adverse events will include all events reported during the study per the treatment emergent definition. Adverse Events will be coded using the Medical Dictionary for Regulatory Activities.

The following presentations will be shown:

- an overall summary showing, the number and percentage of subjects with an adverse event, a mild, moderate or severe event, an event related to study drug, an event related to lumbar puncture procedure, a serious event, a serious event related to study drug, an event leading to study drug withdrawal, an event leading to study withdrawal, and a fatal event.
- incidence and frequency by system organ class and preferred term
- incidence and frequency by preferred term
- incidence and frequency, by preferred term, at least 5% of subjects in either study group
- incidence of adverse events by maximum severity by preferred term
- incidence and frequency of severe events by system organ class and preferred term
- incidence and frequency of events related to study drug by system organ class and preferred term
- incidence and frequency of events related to lumbar puncture (as assessed by the investigator) by system organ class and preferred term

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

- incidence by system organ class and preferred term that occurred within 24 hours of dosing procedure
- incidence by system organ class and preferred term that occurred within 72 hours of dosing procedure
- incidence and frequency of serious adverse events by system organ class and preferred term
- incidence and frequency of serious adverse events related to study drug by system organ class and preferred term
- incidence of events that led to fatal outcome
- incidence of events leading to study drug withdrawal by system organ class and preferred term
- incidence and frequency of adverse events over time by system organ class and preferred term.

- incidence of events leading to withdrawal from study by system organ class and preferred term
- incidence and frequency of adverse events antibody status by system organ class and preferred term.

A separate table will also be presented to show an overall summary of COVID-19 pandemic related AEs.

Listings of the following events will be produced.

- AEs
- SAEs
- AE with fatal outcome
- AEs led to study drug withdrawal
- AEs led to withdrawal from study
- AEs related to lumbar puncture procedure
- AEs related to study drug

5.7.2. Clinical laboratory data

Laboratory data will be evaluated to determine the incidence of abnormalities that emerge during the course of the study. Changes in laboratory evaluations will be presented relative to baseline, which is defined as the closest non-missing visit prior to the first dose.

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

The following clinical laboratory parameters are to be assessed:

- Hematology panel: complete blood count with differential and platelet count (hematocrit, hemoglobin, platelets, red blood cell count [RBC], white blood cell count [WBC], basophils, eosinophils, lymphocytes, monocytes, neutrophils)
- Blood chemistry panel: albumin, total bilirubin, direct and indirect bilirubin, alkaline phosphatase, alanine aminotransferase (ALT) (SGPT), aspartate aminotransferase (AST) (SGOT), gamma-glutamyl transferase (GGT), sodium, potassium, calcium, chloride, phosphate, blood urea nitrogen (BUN), creatinine, cystatin, creatine phosphokinase, creatine kinase, bicarbonate (CO2), glucose, total protein.
- Urinalysis: specific gravity, pH, protein, glucose, ketones, bilirubin, blood, red blood cells, white blood cells, epithelial cells, bacteria, casts and crystals
- Urine total protein assessed by local laboratories
- CSF analysis: RBC, WBC, protein, glucose.
- Coagulation: PT (prothrombin time), aPTT (activated partial thromboplastin time), and INR (international normalized ratio).

Each hematology, blood chemistry, coagulation and CSF laboratory parameter will be flagged as "low", "normal" or "high" relative to the parameter's normal range or as "unknown" if no result is available.

For each urinalysis laboratory parameter, the number and percentage of subjects experiencing postdose shifts to abnormal will be summarized.

For each hematology, blood chemistry, coagulation and CSF parameter, the number and percentage of subjects experiencing postdose shifts to 'low' or 'high' will be summarized. In each summary, the denominator for the percentage is the number of subjects at risk for the shift. The number at risk for the shift to low is the number of subjects whose baseline value was not low and who had at least one post baseline value. The number at risk for the shift to high is the number of subjects whose baseline value was not high and who had at least one-post baseline value. Subjects will be counted only once for each parameter and each shift regardless of how many postdosing assessments had that type of shift. All post baseline data will be used in the shift tables, regardless of whether a scheduled or unscheduled visit.

Also, for each parameter, the incidence of shift to low will be summarized using the minimum post baseline values. Shift to low includes subjects with a normal, high, or unknown baseline value and at least one post baseline value of the given test. Similarly, the incidence of shift to high will be summarized using the maximum post baseline values. Shift to high includes subjects with a low, normal, or unknown baseline value and at least one post baseline value.

To evaluate potential serious hepatotoxicity subjects with a post baseline AST and/or ALT value ≥3 times the upper limit of normal (ULN) and a post baseline bilirubin value >2 times ULN at any time, not necessarily concurrent, will be listed with their values. In addition, a plot will be

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

presented with each subject's maximum post baseline AST or ALT value relative to the ULN against the subject's maximum post baseline bilirubin value relative to the ULN; values do not have to be concurrent. All post baseline data will be used to identify abnormalities, regardless of whether a scheduled or unscheduled visit.

Summary statistics for actual values and change from baseline in laboratory values will be summarized by study group and visit. Line plots for chemistry, hematology and CSF showing median value for each study group at each visit will also be presented.

For a parameter if the local and central result are available with the same date and time then only the central analysis result will be considered for presentations by visit. In a situation where two or more results have the same date and time and are both central (or both local) then we will check with the Safety physician if it would be more appropriate or take the highest or lowest value.

Listings of all chemistry, hematology, coagulation, CSF, urinalysis values and pregnancy test (serum and urine) will be provided.

For platelets a summary of the proportion of participants with a post baseline platelet count below the lower limit of normal on at least 2 consecutive measurements will be presented.

The urine protein positive (trace, +1, +2, +3, +4 or >0.2g/L) results will be presented, number of subjects with positive result at baseline, with a negative result at baseline and at least one and at least two positive post baseline result respectively.

5.7.3. Neurological Examinations

5.7.3.1. Hammersmith infant Neurological Examination (Sections 1 and 3) (HINE 1 and 3) in infantile-onset SMA population

Sections 1 (neurological items) and 3 (behavior) of the HINE serve as the basis for neurological examinations for infantile-onset SMA subjects < 2 years old in cohort 1. These are to be assessed at Screening, predose and at 3, 6, and 24 hours postdose on Day 1, predose and 1, 3 hours postdose on Days 121 and 241, and on Day 302/ET

The neurological items comprise cranial nerve function (facial appearance, eye appearance, auditory response, visual response, sucking/swallowing), posture (head in sitting position, trunk in sitting position, arms at rest, hands, legs in sitting position, legs in supine and standing positions, feet in supine and standing positions), movements (quantity, quality), tone (scarf sign, passive shoulder elevation, pronation/supination, adductors, popliteal angle, ankle dorsiflexion, pulled to sit, ventral suspension), and reflexes and reactions (tendon reflexes, arm protection, vertical suspension, lateral tilting, forward parachute). Behavior is comprised of state of consciousness, emotional state, and social orientation.

Since the number of patients below 2 years of age is expected to be small the results of HINE 1 and 3 testing will be listed only.

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

5.7.3.2. Neurological Examination in patients 2 years or older

Neurological examinations include assessment of mental status, level of consciousness, cranial nerves, reflexes, motor system, coordination/cerebellar function and sensation – temperature and vibration. These are to be assessed at Screening, predose and at 3, 6, and 24 hours postdose on Day 1, predose and at 1, 3 hours on Days 121 and 241, and on Day 302/ET.

The results collected for the majority of the tests are classified as 'normal' or 'abnormal', however the assessment of sensations is reported as 'present' or 'absent' and the assessment of reflexes is captured on an ordinal scale. For each test it is recorded if secondary to SMA.

Acute Effects After Dosing: **Reflex assessments**

On each dosing day, each predose reflex will be compared with all the post dosing assessments with 24 hours to see if the subject has changed (gained or lost) or maintained, reflexes. For this comparison the predose reflex status will be categorized into two categories: 0 and >=1. For each reflex, visit, predose category the number and proportion of subjects who change or maintained will be presented.

Acute Effects After Dosing: Other assessments with binary response (normal/abnormal; present/absent)

On each dosing day, the shifts from predose assessment on that day to postdose (i.e., from predose to 3, 6 and 24 hours postdose) will be determined. For each test and post dosing time point, the number and proportion of subjects who moved from 'normal' to 'abnormal' will be presented, for sensations the number and proportion who move from 'present' to 'absent' will be presented.

Chronic effects: Other assessments with binary response (normal/abnormal; present/absent)

For each test the shift from baseline to (the predose value on Day 1) to the predose value on later dosing days and to Day 1, 121, 241 and on Day 302/ET will be determined, the number and proportion of subjects who moved from 'normal' to 'abnormal' will be presented, for sensations the number and proportion who move from 'present' to 'absent' will be presented.

5.7.4. ECG

ECGs are to be recorded at Days 1, 121, 241, and 302/ET, predose (Day 1) and post dosing (Days 1, 121, 241) at various time points on dosing days and end predose at 302/ET. The ECGs are assessed at a central reading laboratory and the results provided as external vendor data. On the eCRF the investigator interpretation is collected as normal, abnormal and abnormal AE and the investigator determination will be used in analyses.

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

The ECG test includes heart rate, QT interval, QTcF interval and RR interval. Summary statistics for actual values and change from baseline in each ECG parameter will be presented by study group and visit.

ECGs will be analyzed using two approaches.

Qualitative analysis

The determination of whether an ECG is abnormal not AE, abnormal AE made by the Investigator will be used for this qualitative analysis.

The number and percentage of subjects with shifts from normal to each of the categorical values denoting an abnormal scan (abnormal not AE, abnormal AE) will be summarized by study group. All post baseline data will be used to identify abnormalities, regardless of whether a scheduled or unscheduled visit.

Corrected Fridericia QT interval QTcF abnormalities (outliers)

QTcF will be examined to determine the incidence of clinically relevant abnormalities. The number of subjects evaluated and the number of subjects with clinically relevant abnormalities will be presented.

The first criteria for clinically relevant post dosing abnormalities (secondary safety endpoint) that will be considered is:

- Post baseline QTcF of > 500 msec and Maximum increase from baseline to post baseline QTcF >60 msec

Other criteria are:

- Maximum increase from baseline QTcF > 30 to 60 msec
- Maximum increase from baseline QTcF > 60 msec
- Maximum post baseline QTcF > 480 to 500 msec
- Maximum post baseline QTcF > 500 msec

ECG summaries will be presented for all subjects and for subjects with post baseline. Listings will include both the central and eCRF information.

5.7.5. Vital Signs

The analysis of vital signs will be approached in two ways.

Vital signs are to be measured at Screening, predose and at 1, 2, 4, 6, 8 and 24 hours postdose on Day 1, predose and at 1 and 6 hours postdose on Days 121 and 241 At each of these times, temperature, heart rate, respiratory rate, systolic and diastolic blood pressure, and pulse oximetry will be measured.

Acute Effects After Dosing

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

On each dosing day, the change in each vital sign from predose on that day to postdose (i.e., from predose to 1 hour postdose, to 2 hours postdose, etc.) will be calculated. Summaries of actual values and change from predose for each dosing day will be presented.

Chronic Effects

To examine for possible chronic effects, the change from baseline (the predose value on Day 1) to the predose value on later dosing days and to Day 302/ET will be determined. Summaries of actual values and changes from baseline for subsequent dosing days (Days 121 and 241) and for Day 302/ET will be presented.

Vital signs (temperature, pulse, systolic and diastolic blood pressure and pulse oximetry) will also be examined to determine the incidence of potentially clinically relevant abnormalities.

Shift analysis

The number of subjects evaluated and the number of subjects with potentially clinically relevant abnormalities will be presented. All post baseline data will be used to identify abnormalities, regardless of whether a scheduled or unscheduled visit. The criteria for potentially clinically relevant postdose abnormalities are shown in Table 6 below

Table 6: Criteria to determine potentially clinically relevant abnormalities in vital signs

Vital Sign	Criteria for Abnormalities
Temperature	<36°C >38°C
Pulse	<60 bpm >100 bpm
Systolic Blood Pressure	<90 mmHg >140 mmHg >160 mmHg
Diastolic Blood Pressure	<50 mmHg >90 mmHg >100 mmHg
Weight	7% or more decrease from baseline 7% or more increase from baseline
Respiratory Rate	<12 breaths/min

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

	>20 breaths/min

5.7.6. Growth parameters

Growth parameters for infantile-onset and later-onset SMA population include weight, height/length ulnar length, length for age (<18 years of age) and weight for age (<18 years of age).

The growth parameters are to be assessed during screening, on dosing days (Days 1, 121, and 241) and on Day 302/ET.

The 2000 CDC Growth Charts (ages 2 to <20 years) will be used to assess the weight change for older subjects. The National Center for Health Statistics provides a SAS macro which can be downloaded from their website [https://www.cdc.gov/growthcharts/clinical_ charts.htm] and this will be utilized to calculate the weight for age percentiles for each participant in the later onset population.
Subjects will be cross referenced with these files, given the age and sex of the subject to determine below which percentile they lie for each parameter.

For patients <18 years of age the weight and length for age percentiles will be determined and summarized over time.

Summary statistics for actual values and change from baseline in each growth parameter will be presented by visit.

Baseline growth parameters for subjects <18 years of age at baseline, will be summarized and also the frequency and percentage for weight and height for age will also be presented for each of in the following categories: ≤ 1st, ≤ 3rd, ≤5th, ≤15th, ≤25th, ≤50th and > 50th percentile.

[redacted]

5.7.8. LP opening pressure

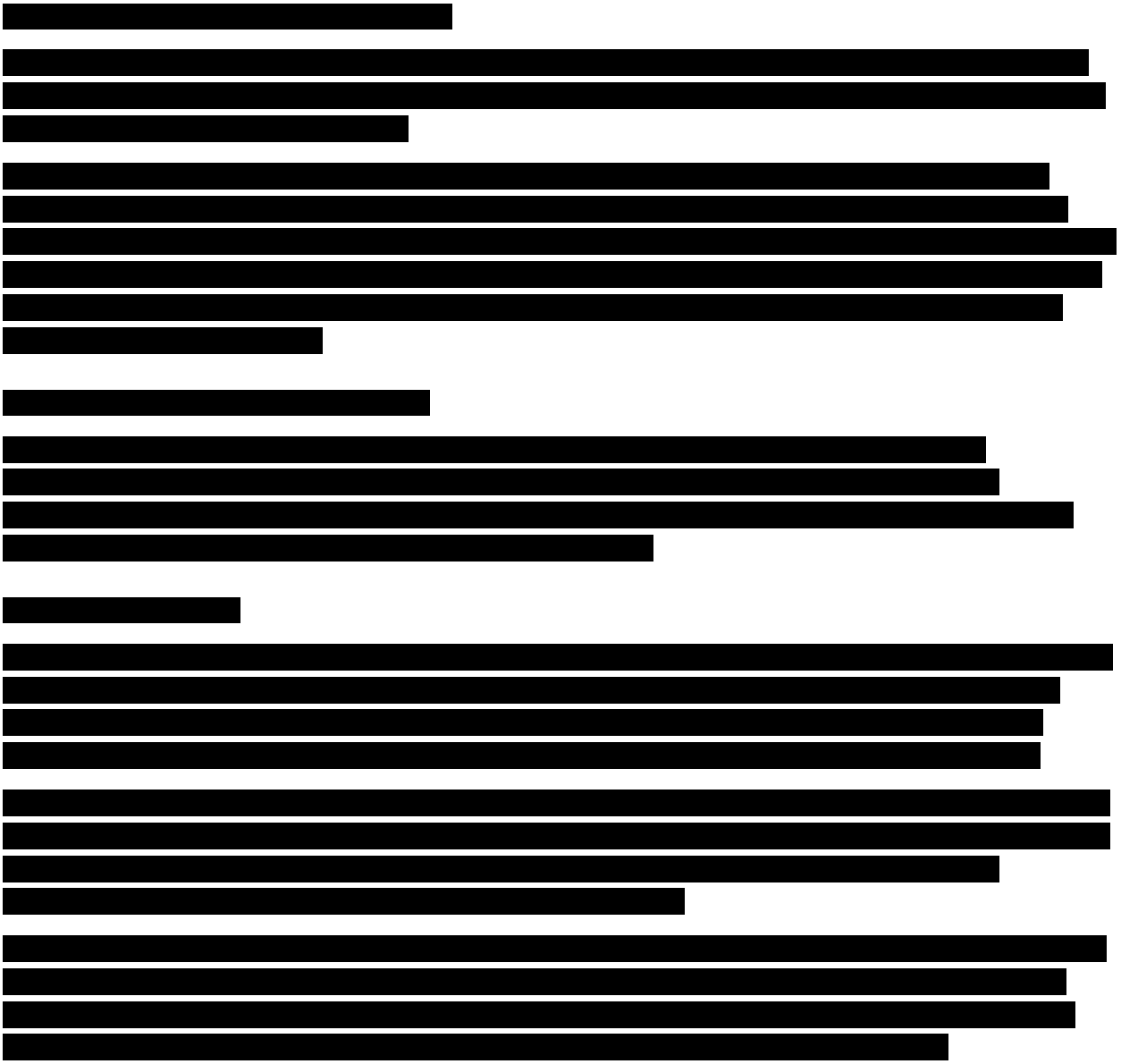
LP opening pressure will only be measured at dosing days (Days 1, 121, and 241). A lack of published data exists in terms of the variability of opening pressure between and within subjects. It is suspected that a simple change in position of a subject during the measurement or change of instrument used contributes to variability in the measurement.

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

The actual value and change from baseline to each visit will be summarized, all the data collected will be listed.

5.7.9. X-ray Examination of Spine

The X-rays will be used to determine the severity of scoliosis by measuring the Cobb angle of later onset patients in Cohort 1 and all patients in Cohort 2. The spine X-ray is performed at screening with the participant in a sitting or supported sitting position following an image acquisition guideline. Summary and listing of baseline Cobb angle will be presented.



DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

[redacted]

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

- [redacted]
- [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

[redacted]

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

[redacted]

- [redacted]
- [redacted]

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

5.8. Other Analyses

Not applicable.

5.9. Statistical Considerations for Interim Analysis

Not applicable

6. Changes from Protocol-Specified Analyses

[redacted]

7. Summary of Changes from the Previous Version of the SAP

8. References

[redacted]

[redacted]

Bohdana Ratitch, Ilya Lipkovich, Michael O'Kelly. Combining Analysis Results from Multiply Imputed Categorical Data. PharmaSUG 2013 - Paper SP03.

De Sanctis R, Coratti G, Pasternak A, et al. Developmental milestones in type I spinal muscular atrophy. Neuromuscul Disord. 2016;26(11):754-759. Epub 2016/10/05.

Glanzman AM, Mazzone E, Main M, et al. The Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders (CHOP INTEND): test development and reliability. Neuromuscul Disord. 2010;20(3):155-61.

Glanzman AM, McDermott MP, Montes J, et al. Validation of the Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders (CHOP INTEND). Pediatr Phys Ther. 2011;23(4):322-6.

Guy W. ECDEU Assessment Manual for Psychopharmacology. Rockville, MD: U.S. Department of Health, Education, and Welfare; 1976.

Haataja L, Mercuri E, Regev R, et al. Optimality score for the neurologic examination of the infant at 12 and 18 months of age. J Pediatr. 1999;135(2 Pt 1):153-61.

[redacted]

[redacted]

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

Main M, Kairon H, Mercuri E, et al. The Hammersmith functional motor scale for children with spinal muscular atrophy: a scale to test ability and monitor progress in children with limited ambulation. Eur J Paediatr Neurol. 2003;7(4):155-9.

Matsumoto H, Clayton-Krasinski DA, Klinge SA, et al. Development and initial validation of the assessment of caregiver experience with neuromuscular disease. J Pediatr Orthop. 2011;31(3):284-92.

Mazzone ES, Mayhew A, Montes J, et al. Revised upper limb module for spinal muscular atrophy: Development of a new module. Muscle Nerve. 2016 Epub 2016/10/04.

O'Hagen JM, Glanzman AM, McDermott MP, et al. An expanded version of the Hammersmith Functional Motor Scale for SMA II and III patients. Neuromuscul Disord. 2007;17(9-10):693-7.

Varni JW, Seid M, Rode CA. The PedsQL: measurement model for the pediatric quality of life inventory. Med Care. 1999;37(2):126-39.

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

Wirth B, Brichta L, Schrank B, et al. Mildly affected patients with spinal muscular atrophy are partially protected by an increased SMN2 copy number. Hum Genet. 2006;119(4):422-8.

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

9. APPENDICES

APPENDIX A. Derivation of the dates

i Table 9: Seeds used for each endpoint in MI

Endpoint	Seed
HFMSE	25980495
RULM	47886416

ii Adverse Events

In the situation where worsening in severity (including seriousness) occurs for an adverse event, study sites are instructed to enter an end date and start a new record for the adverse event. In the new record, the changed severity is to be recorded and the start date will be set to be the end date of the previous record. Data linking those records are collected in the data base. Consider three scenarios:

- The first AE record occurs prior to the first dose, the second AE record occurs on or after the first dose, and the severity increases: Only the second record will be counted as treatment-emergent.
- Both records occur on or after the first dose: If the AE severity on the second record is worse than the severity on the first record, then count both records as treatment-emergent.
- Both records occur on or after the first dose: If worsening in severity and and seriousness then count both records as treatment-emergent.

Of note, when counting the total number of treatment-emergent events, events linked through change in severity will still be counted as separate events.

For events with missing start or stop dates, the following criteria will be used for the purpose of identifying treatment-emergent adverse events:

- If both the start and stop date for a particular event are missing, then the event is considered to have occurred on or after the first dose of study treatment.
- If the start date for a particular event is missing and the stop date/time falls after the first dose date/time, then the event is considered to have occurred on or after the first dose of study treatment.
- If the start time is missing and the start date is same as the first dosing date, then the event is considered to have occurred on or after the first dose of study treatment.

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

- If it cannot be determined whether or not an event has occurred on or after dosing due to a missing or partial date, then the event will be assumed to have occurred on or after the first dose for the purpose of identifying treatment-emergent adverse events.

Specifically, let AESTDT denote the start date of an adverse event and TRTSTDT be the start date of study treatment. For the purpose of identifying the treatment emergent adverse events, the following algorithm will be used for the imputation of missing or partial date:

- If AESTDT is completely missing or the year is missing, then impute AESTDT to TRTSTDT.
- If, in AESTDT, year is present and month/day are missing and year is equal to the year portion of TRTSTDT, then impute the month/day portion of AESTDT to the month/day portion of TRTSTDT.
- If, in AESTDT, year is present and month/day are missing and year is not equal to the year portion of TRTSTDT, then impute the month/day portion of AESTDT to January 01.
- Consider the situation in AESTDT where year and month are present with only day missing. If the year and month are the same as those for TRTSTDT, then impute day in AESTDT with day in TRTSTDT. Otherwise, impute the day in AESTDT with the first day of the month.

It is important to emphasize that the imputed date will not be used for calculations such as onset and duration of an adverse event.

iii Concomitant Therapy and Procedure

In order to define concomitant therapies with missing start or stop dates, the following additional criteria will be used:

- If both the start and stop dates of a particular therapy were missing, that therapy is considered concomitant.
- If the start date of a therapy was missing and the stop date of that therapy fell on or after the date of dosing, that therapy is considered concomitant.
- If the start date of a therapy was prior to the date of dosing and the stop date of that therapy was missing and the therapy was listed as continuing, that therapy is considered concomitant.
- If the start date of a therapy was prior to the date of dosing and the stop date of that therapy was missing and the therapy was listed as *not* continuing, that therapy is considered concomitant.
- If the start/stop date of a therapy is partial then where it is not possible to rule out that it was not taken concomitantly it will be considered concomitant.

DocuSign Envelope ID: 601DBC7A-2C2F-44D5-97AD-76A7B22B7044

Denote the end date of medications as CMENDT and the study treatment start date as TRTSTDT. The medication is classified concomitant provided any of the following is NOT true:

- CMENDT is complete and CMENDT is less than TRTSTDT.
- Day of CMENDT missing and year/month of CMENDT is strictly before year-month of TRTSTDT.
- Month of CMENDT is missing and year of CMENDT is strictly before year of TRTSTDT.

DocuSign

Certificate Of Completion

Envelope Id: 601DBC7A2C2F44D597AD76A7B22B7044
Status: Completed
Subject: Complete with DocuSign: 232SM203 Part C SAP final.docx
Vendor Name:
Source Envelope:
Document Pages: 51
Signatures: 3
Envelope Originator:
Certificate Pages: 2
Initials: 0
[redacted]
AutoNav: Enabled
225 Binney St
EnvelopeId Stamping: Enabled
Cambridge, MA 02142
Time Zone: (UTC-05:00) Eastern Time (US & Canada)
[redacted]
IP Address: [redacted]

Record Tracking

Status: Original
11 October 2023 | 10:59
Holder: [redacted]
[redacted]
Location: DocuSign

Signer Events	Signature	Timestamp
[redacted] Principal Biostatistician Biogen Security Level: Email, Account Authentication (None)	DocuSigned by: [redacted] Signature Adoption: Pre-selected Style Using IP Address: [redacted]	Sent: 11 October 2023 | 11:02 Viewed: 11 October 2023 | 11:54 Signed: 11 October 2023 | 11:54
Electronic Record and Signature Disclosure: Not Offered via DocuSign		
[redacted], Spinal Muscular Atrophy (SMA) Biogen Security Level: Email, Account Authentication (None)	DocuSigned by: [redacted] Signature Adoption: Pre-selected Style Using IP Address: [redacted]	Sent: 11 October 2023 | 11:02 Viewed: 11 October 2023 | 11:17 Signed: 11 October 2023 | 11:17
Electronic Record and Signature Disclosure: Not Offered via DocuSign		
[redacted] Biogen Security Level: Email, Account Authentication (None)	DocuSigned by: [redacted] Signature Adoption: Pre-selected Style Using IP Address: [redacted]	Sent: 11 October 2023 | 11:02 Viewed: 11 October 2023 | 11:02 Signed: 11 October 2023 | 11:08
Electronic Record and Signature Disclosure: Not Offered via DocuSign		

In Person Signer Events	Signature	Timestamp

Editor Delivery Events	Status	Timestamp

Agent Delivery Events	Status	Timestamp

Intermediary Delivery Events	Status	Timestamp

Certified Delivery Events	Status	Timestamp

Carbon Copy Events	Status	Timestamp

Witness Events	Signature	Timestamp

Notary Events	Signature	Timestamp

Envelope Summary Events	Status	Timestamps
Envelope Sent	Hashed/Encrypted	11 October 2023 | 11:02
Certified Delivered	Security Checked	11 October 2023 | 11:02
Signing Complete	Security Checked	11 October 2023 | 11:08
Completed	Security Checked	11 October 2023 | 11:54

Payment Events	Status	Timestamps

Signature Page

Document Name: 232SM203 Part C Statistical Analysis Plan

Document Title: Escalating Dose and Randomized, Controlled Study of Nusinersen (BIIB058) in Participants With Spinal Muscular Atrophy

Signed by	Role	Date / Time (UTC)
[redacted]	Signing as Approver	12-Oct-2023 14:34:45

Signature Page

Document Name:

Document Title:

Signed by	Role	Date / Time (UTC)