22 December 2025

Clinicaltrials.gov Results Posting

Re: *SRK-015-003 Results | Statistical Analysis Plan*

Sponsor:	Scholar Rock, Inc. 301 Binney Street, 3rd Floor Cambridge, MA 02142 USA
Protocol Number:	SRK-015-003 (SAPPHIRE)
Protocol Title:	Phase 3, Double-Blind, Placebo-Controlled Trial to Evaluate the Efficacy and Safety of Apitegromab (SRK-015) in Patients with Later-Onset Spinal Muscular Atrophy Receiving Background Nusinersen or Risdiplam Therapy
NCT Number:	05156320
Protocol Document Date:	11 June 2024
Statistical Analysis Plan Document Date:	30 August 2024

Thank you,
Regulatory Affairs Record Administrator
Scholar Rock, Inc.
clinicaltrials@scholarrock.com

Scholar Rock
301 Binney Street, 3rd Floor
Cambridge, MA 02142

(857) 259-3860
scholarrock.com

Statistical Analysis Plan

Protocol Number:	SRK-015-003
Protocol Title:	Phase 3, Double-Blind, Placebo-Controlled Trial to Evaluate the Efficacy and Safety of Apitegromab (SRK-015) in Patients with Later-Onset Spinal Muscular Atrophy Receiving Background Nusinersen or Risdiplam Therapy
Protocol Version:	Version 3.0
Protocol Date:	11JUN2024
Trial Name:	SAPPHIRE
SAP Version:	**V1.0**
SAP Date:	**30AUG2024**
Author:	[redacted]
Sponsor:	Scholar Rock, Inc. 301 Binney Street, 3rd Floor Cambridge, MA 02142

Notice of Confidential and Proprietary Information:

REVISION HISTORY

Version #	Date	Revision Summary
1.0	30AUG2024	Original Version

APPROVAL SIGNATURE

[redacted]

[redacted] Date

[redacted]

[redacted] Date

[redacted]

[redacted] Date

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

Abbreviation	Description
ACEND	Assessment of Caregiver Experience with Neuromuscular Disease
ADA	antidrug antibody
AE	adverse event
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
aPTT	activated partial thromboplastin time
AR	autoregression
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
BILI	bilirubin
BLQ	below the level of quantification
BMI	body mass index
BUN	blood urea nitrogen
C-SSRS	Columbia-Suicide Severity Rating Scale
CI	confidence interval
cm	centimeters
CMH	Cochran-Mantel-Haenszel
COVID-19	coronavirus disease 2019
CSR	clinical study report
CTCAE	Common Terminology Criteria for Adverse Events
ECG	electrocardiogram
eCRF	electronic case report form
EOS	end of study
EOT	end of treatment
ET	early termination
EXP	Exploratory Subpopulation
GGT	gamma-glutamyl transferase
hCG	human chorionic gonadotropin
HFMSE	Hammersmith Functional Motor Scale - Expanded
ICE	intercurrent event
ICF	informed consent form

Abbreviation	Description
ICH	International Council for Harmonisation
IDMC	Independent Data Monitoring Committee
IEC	Independent Ethics Committee
ITT	Intention-to-Treat
IV	intravenous
IWRS	interactive web-based randomization system
LDH	lactate dehydrogenase
LS	least squares
MAR	missing-at-random
MCH	mean corpuscular hemoglobin
MCMC	Markov Chain Monte Carlo
MedDRA	Medical Dictionary for Regulatory Activities
MEP	Main Efficacy Population
MGRS	Multicentre Growth Reference Study
MI	multiple imputation
MITT	Modified Intention-to-Treat
MMRM	mixed effects model with repeated measurement
MNAR	missing-not-at-random
N/A	not applicable
NCI	National Cancer Institute
NCS	not clinically significant
PD	pharmacodynamic(s)
PEDI-CAT	Pediatric Evaluation of Disability Inventory Computer Adaptive Test
PK	pharmacokinetic(s)
PMM	pattern-mixture model
PP	Per-Protocol
PT	preferred term
PT/INR	prothrombin time/international normalized ratio
PROMIS	Patient-Reported Outcomes Measurement Information System
QC	quality control
QTc interval	corrected QT interval
QTcB interval	QT interval corrected by Bazett's formula

Abbreviation	Description
QTcF interval	QT interval corrected by Fridericia's formula
RBC	red blood cell
REML	restricted maximum likelihood
RULM	Revised Upper Limb Module
SAE	serious adverse event
SAP	statistical analysis plan
SAS®	Statistical Analysis System
SD	standard deviation
SMA	spinal muscular atrophy
SMN	survival motor neuron
SMQ	Standardized MedDRA Queries
SoA	schedule of activities/assessments
SOC	System Organ Class
SOP	standard operating procedure
TEAE	treatment-emergent adverse event
TFLs	tables, figures, and listings
ULN	upper limit of normal
V14	Visit 14
WBC	white blood cell
WHO	World Health Organization

1. INTRODUCTION

The intent of this document is to provide guidance for analyses of data for the Phase 3 clinical trial of apitegromab in patients with spinal muscular atrophy (SMA). This statistical analysis plan (SAP) has been developed based on the current version of the clinical study protocol SRK-015-003 (SAPPHIRE; Version 3.0, 11JUN2024). The SAP will be consistent with the latest protocol amendment where applicable.

The purpose of this SAP is to ensure that the data listings, summary tables, and figures that will be produced, as well as the statistical methodologies that will be used, are complete and appropriate to allow valid conclusions regarding the study objectives. This SAP follows the International Council for Harmonisation (ICH) E9 guidance on statistical principles for clinical trials.

1.1. Scope of Work

This SAP covers the planned analyses of the Phase 3 trial (SRK-015-003 [SAPPHIRE]).

All analyses of efficacy, safety, pharmacokinetic (PK), pharmacodynamic (PD), and antidrug antibody (ADA) endpoints described in this plan will be conducted in the final analysis.

Unblinding will occur and safety and efficacy analyses will be conducted when all patients complete the Treatment Period (12-month assessment at Visit 14 [V14]), at which time all patients will have completed study drug administrations. If there are patients still being followed in the Safety Follow-up Period, during which patients will be followed for safety but will no longer receive study drug, data collected after all patients completed the Treatment Period will be provided where applicable.

The SAP will be finalized prior to the database lock or unblinding of treatment assignment, whichever occurs first. No changes to the SAP after finalization are expected. However, if any changes are necessary, the SAP will be updated to a new version, with all revisions documented.

1.2. Timings of Analyses

During the study conduct, an Independent Data Monitoring Committee (IDMC) will periodically review the safety data per the IDMC Charter. The description of the IDMC analyses and meeting requirements can be found in the IDMC Charter.

1.3. Tables, Figures, and Listings

A detailed description of the planned Tables, Figures, and Listings (TFLs) to be presented in the clinical study report (CSR) will be provided in the TFL shells document. When the SAP and TFL shells are agreed upon and finalized, they will serve as the main source for data analyses for the CSR.

2. STUDY OBJECTIVES AND ENDPOINTS

2.1. Main Efficacy Population: Efficacy and Additional Objectives/Endpoints

The MEP consists of patients with later-onset SMA who are 2 through 12 years old (inclusive) at Screening. The primary efficacy, key secondary efficacy, and additional objectives evaluated on the MEP are listed in Table 1.

Table 1: Main Efficacy Population: Efficacy and Additional Objectives/Endpoints

Objectives	Endpoints
Primary Efficacy	
Assess the efficacy of apitegromab compared with placebo using the HFMSE in patients 2 through 12 years old	Change from Baseline in HFMSE total score at 12 months
Key Secondary Efficacy	
Assess the efficacy of apitegromab compared with placebo by measuring changes in upper limb function using the RULM in patients 2 through 12 years old	Change from Baseline in RULM total score at 12 months
Assess the efficacy of apitegromab compared with placebo based on the number of patients with clinical improvement in patients 2 through 12 years old	Proportion of patients with ≥3-point change from Baseline in the HFMSE total score at 12 months
Assess the efficacy of apitegromab compared with placebo by measuring changes in number of WHO motor development milestones in patients 2 through 12 years old	Change from Baseline in number of WHO motor development milestones attained at 12 months
Other Secondary Efficacy	
Further assess the efficacy of apitegromab compared with placebo by evaluating changes in additional motor function outcome measures and changes in HFMSE at other prespecified time points in patients 2 through 12 years old	• Proportion of patients achieving various magnitudes of change in HFMSE score from Baseline at 12 months • Proportion of patients achieving various magnitudes of change in RULM score from Baseline at 12 months • Proportion of patients who attain a new WHO motor development milestone relative to Baseline at 12 months • Change from Baseline in HFMSE total score at other prespecified time points

Objectives	Endpoints
	• Change from Baseline in RULM total score at other prespecified time points • Change from Baseline in number of WHO motor development milestones attained at other prespecified time points
Additional Efficacy	
Assess the efficacy of apitegromab compared with placebo by measuring changes from Baseline in motor function across the Treatment Period using the HFMSE in patients 2 through 12 years old	Change from Baseline in HFMSE total score across time during the 12-month Treatment Period
Assess the time to therapeutic effect of apitegromab compared with placebo using the HFMSE in patients 2 through 12 years old	Time to therapeutic effect (≥3-point change from Baseline in HFMSE total score) compared between apitegromab and placebo
Additional Other	
Evaluate the effects of apitegromab on patient/caregiver-reported disability, fatigability, and suicidal ideation and behavior in patients 2 through 12 years old	• Change from Baseline in PEDI-CAT • Change from Baseline in PROMIS Fatigue Questionnaire • Change from Baseline in ACEND • Change from Baseline in C-SSRS

Abbreviations: ACEND, Assessment of Caregiver Experience with Neuromuscular Disease; C-SSRS, Columbia-Suicide Severity Rating Scale; HFMSE, Hammersmith Functional Motor Scale - Expanded; PEDI-CAT, Pediatric Evaluation of Disability Inventory Computer Adaptive Test; PROMIS, Patient-Reported Outcomes Measurement Information System; RULM, Revised Upper Limb Module; WHO, World Health Organization.

2.1.1. Motor Function Outcome Measures

2.1.1.1. Hammersmith Functional Motor Scale - Expanded

The Hammersmith Functional Motor Scale - Expanded (HFMSE), which is validated in SMA, assesses the physical abilities of patients with Type 2 and Type 3 SMA (O'Hagen 2007, Glanzman 2011). It consists of 33 items graded on a scale of 0 to 2, where 0 denotes unable, 1 denotes performed with modification or adaptation, and 2 denotes performed without modification or adaptation. The item scores are summed to give a total score with a maximum of 66. In conducting the HFMSE test, Item 3 (ie, one hand to head in sitting) will evaluate both hands in which the higher score will be recorded and used in calculating the total score. A higher total score indicates better motor function performance.

2.1.1.2. Revised Upper Limb Module

The Revised Upper Limb Module (RULM), which is validated in SMA, is a 19 scorable-item assessment of upper limb function in nonambulatory patients with SMA [young children as well as adults] (Mazzone 2017). The 19 scorable items test functions that relate to everyday life, such as bringing hands from the lap, pressing a button, and picking up a token. With the exception of 1 item (ie, open Ziploc® container) with a binary score (can/cannot score), 18 of the 19 scorable items are scored 0, 1, and 2, where 0 denotes unable, 1 denotes able with modification, and

2 denotes able with no difficulty. The maximum score achievable is 37. To calculate the RULM total score, for items where scores are collected for both hands, only the higher of the 2 scores will be used to calculate the total score. The RULM will be completed by patients who are ≥30 months old at the time of the Baseline assessment. A higher total score indicates better motor function performance.

2.1.1.3. WHO Motor Development Milestones

The World Health Organization (WHO) motor development milestones are a set of 6 distinct gross motor milestones that are considered to be universal and fundamental to acquiring the ability to walk independently (Wijnhoven 2004). The WHO Multicentre Growth Reference Study (MGRS) performance criteria, described in Table 2, are being used to assess each of the 6 motor milestones. The motor development milestones of patients are assessed based on the age windows for achieving these milestones, which were validated in an international study (WHO Multicentre Growth Reference Study Group 2006). A higher WHO motor development milestones value indicates more milestones are achieved by a patient.

Table 2: WHO Multicentre Growth Reference Study Performance Criteria

Gross Motor Milestone	MGRS Performance Criteria
Sitting without support	Child sits up straight with the head erect for at least 10 seconds. Child does not use arms or hands to balance body or support position.
hands and knees crawling	Child alternately moves forward or backward on hands and knees. The stomach does not touch the supporting surface. There are continuous and consecutive movements, at least 3 in a row.
Standing with assistance	Child stands in upright position on both feet, holding onto a stable object (eg, furniture) with both hands without leaning on it. The body does not touch the stable object, and the legs support most of the body weight. Child thus stands with assistance for at least 10 seconds.
Walking with assistance	Child is in upright position with the back straight. Child makes sideways or forward steps by holding onto a stable object (eg, furniture) with one or both hands. One leg moves forward while the other supports part of the body weight. Child takes at least 5 steps in this manner.
Standing alone	Child stands in upright position on both feet (not on the toes) with the back straight. The legs support 100% of the child's weight. There is no contact with a person or object. Child stands alone for at least 10 seconds.
Walking alone	Child takes at least 5 steps independently in upright position with the back straight. One leg moves forward while the other supports most of the body weight. There is no contact with a person or object.

Abbreviations: MGRS, Multicentre Growth Reference Study; WHO, World Health Organization.

2.1.2. Assessment of Patient/Caregiver-Reported Outcomes

2.1.2.1. Pediatric Evaluation of Disability Inventory Computer Adaptive Test

The Pediatric Evaluation of Disability Inventory Computer Adaptive Test (PEDI-CAT) is a questionnaire completed by the caregiver that assesses the patient's ability to perform daily functions (Haley 2005). The same caregiver must fill out the assessment throughout the trial duration. The answers are scored on a 4-point scale ("unable" to "easy"). The test is suitable to assess function in newborns through 21years old; this questionnaire should be completed throughout the duration of the trial (regardless of age). Properties of the PEDI-CAT have been studied in the SMA population. A Rasch analysis with results published in 2016 revealed that the distribution of abilities for the Mobility and Daily Activities domains of the PEDI-CAT is best represented in the Type 2 and Type 3 populations (Pasternak 2016). As the PEDI-CAT is an age-dependent questionnaire, a patient's full date of birth may be required for accurate assessment per local regulations. The domains to be evaluated for this study include daily activities and mobility where standardized scores provided by the PEDI-CAT system will be assessed. A higher score indicates better ability to perform the daily functions or better mobility.

2.1.2.2. Patient-Reported Outcomes Measurement Information System Fatigue Questionnaire

The Patient-Reported Outcomes Measurement Information System (PROMIS) is a person-centered measure intended to be completed by the patient or parent proxy without help from anyone (Ader 2007). The fatigue profile domain measures a range of symptoms, from mild patientive feelings of tiredness to an overwhelming, debilitating, and sustained sense of exhaustion. The PROMIS will be completed by/for patients who are or will turn 5 years old or older at the time of the Baseline assessment; the same questionnaire used at Baseline should be used throughout the duration of the trial (regardless of age). The self-reported PROMIS measures are suitable for children 8 to 17 years old, and the parent proxy-reported PROMIS measures are suited for children 5 to 17 years old. If a caregiver completes this form, the same caregiver must complete the form throughout the trial duration. Patients who are 18 through 21 years old at Screening will complete an adult form of PROMIS. A higher score from the same questionnaire indicates higher degree of tiredness.

2.1.2.3. Assessment of Caregiver Experience with Neuromuscular Disease

The Assessment of Caregiver Experience with Neuromuscular Disease (ACEND) is a validated self-administered instrument for assessing caregiver impact on parents raising children severely affected by neuromuscular disease (Matsumoto 2011). The ACEND is completed by the caregiver. The same caregiver must complete the assessment throughout the trial duration.

The ACEND instrument includes 2 domains, 7 subdomains, and 41 items:

- Domain 1, which examines physical impact, includes 4 subdomains: feeding/grooming/dressing (6 items), sitting/play (5 items), transfers (5 items), and mobility (7 items).
- Domain 2, which examines general caregiver impact, includes 3 subdomains: time (4 items), emotion (9 items), and finance (5 items).

A score for each item is generated based on the 6- and 5-point ordinal scales and is designed that caregivers experiencing less intense caregiving impact received higher scores.

2.1.2.4. Columbia-Suicide Severity Rating Scale

The Columbia-Suicide Severity Rating Scale (C-SSRS) "Baseline/Screening" and "since last visit" version will be used for patients ≥6 years old. The children's versions of the C-SSRS "Baseline/Screening" and "since last visit" will be used for patients who are 4 to 5 years old at Screening until they turn 6 years old.

The C-SSRS involves a series of probing questions to inquire about possible suicidal thinking and behavior, but the full interview is needed only if the initial Screening questions about suicidal ideation and behavior are positive.

There are 5 questions, listed below, in the section of "suicidal ideation", each with a binary response (yes or no). Questions 1 and 2 are the initial Screening questions. Questions 3, 4, and 5 in this section will only be evaluated when the answer to Question 2 is "yes". The positive suicidal ideation is defined as a "yes" answer to any of these 5 questions. Severity of ideation is rated on a 5-point ordinal scale ranging from 1 "Wish to be dead" to 5 "Suicidal intent with plan"; it will be assigned the maximum suicidal ideation category listed below, and a score of 0 will be assigned if no ideation is present. Treatment-emergent suicidal ideation is defined as Baseline suicidal ideation is not present but is positive at any postbaseline visit (Nilsson 2013).

- Question/Category 1: Wish to be Dead
- Question/Category 2: Non-Specific Active Suicidal Thoughts
- Question/Category 3: Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act
- Question/Category 4: Active Suicidal Ideation with Some Intent to Act, without Specific Plan
- Question/Category 5: Active Suicidal Ideation with Specific Plan and Intent

The intensity of ideation section will only be evaluated when the answer to Question 1 and/or 2 is "yes" in the section of "suicidal ideation". The intensity of ideation includes 5 items related to the most severe ideation (eg, frequency, duration, controllability, deterrents, and reasons for ideation) in the "Baseline/Screening" and "since last visit" versions and only 1 item (frequency) related to the most severe ideation in the Children's version of the C-SSRS "Baseline/Screening" and "since last visit".

There are 5 items, listed below, in the section of "suicidal behavior," each with a binary response (yes or no). A positive suicidal behavior is defined as a "yes" answer to any of the 5 questions. Treatment-emergent suicidal behavior is defined as Baseline suicidal behavior is not present but is positive at any postbaseline visit.

1. Actual Attempt (non-fatal)
2. Interrupted Attempt
3. Aborted Attempt

4. Preparatory Acts or Behavior
5. Completed Suicide

Strengths of the C-SSRS include that it has been used and validated with adults and adolescents (Gipson 2015, Posner 2011) and has been found to predict short-term suicidal behavior among high-risk adolescents (Conway 2016). Instruments such as the C-SSRS have been used successfully in children and adolescent patients with various psychiatric disorders that do not involve cognitive impairment. Nevertheless, assessing young children also can be challenging because many may not have reached sufficient cognitive maturity to understand the concept of death.

C-SSRS is being performed because the survival motor neuron (SMN) therapy risdiplam is considered to be a central nervous system-active intervention.

2.2. Main Efficacy Population/Exploratory Subpopulation Combined: Secondary Objectives/Endpoints

The Exploratory Subpopulation (EXP) consists of patients with later-onset SMA who are 13 through 21 years old (inclusive) at Screening. The secondary objectives include safety, tolerability, and PK/PD/ADA evaluated in the Pooled Population (ie, MEP/EXP Combined) and are listed in Table 3.

Table 3: Main Efficacy Population/Exploratory Subpopulation Combined: Secondary Objectives/Endpoints

Objectives	Endpoints
Secondary	
Assess safety and tolerability of apitegromab in all randomized patients with later-onset SMA who receive at least 1 dose of apitegromab	Incidence of TEAEs and SAEs by severity
Characterize the PK of apitegromab in all randomized patients with later-onset SMA who receive at least 1 dose of apitegromab	Apitegromab concentrations in serum from blood samples
Evaluate the PD effects of apitegromab in all randomized patients with later-onset SMA who receive at least 1 dose of apitegromab	Circulating latent myostatin concentrations in blood samples
Evaluate the immunogenicity of apitegromab in all randomized patients with later-onset SMA who receive at least 1 dose of apitegromab	Presence or absence of ADA against apitegromab in serum from blood samples

Abbreviations: ADA, antidrug antibody; PD, pharmacodynamic; PK, pharmacokinetics; SAE, serious adverse event; SMA, spinal muscular atrophy; TEAE, treatment-emergent adverse event.

2.3. Exploratory Subpopulation and Main Efficacy Population/Exploratory Subpopulation Combined: Additional Objectives/Endpoints

Additional objectives evaluated on the EXP and the Pooled Population (ie, MEP/EXP Combined) are listed in Table 4.

Table 4: Exploratory Subpopulation and Main Efficacy Population/Exploratory Subpopulation Combined: Additional Objectives/Endpoints

Objectives	Endpoints
Additional	
Assess the efficacy of apitegromab compared with placebo using the HFMSE in patients 13 through 21 years old and in patients 2 through 21 years old	Change from Baseline in HFMSE total score at 12 months
Assess the efficacy of apitegromab compared with placebo based on the number of patients with clinical improvement or stabilization in patients 13 through 21 years old	Proportion of patients with ≥0-point change from Baseline in the HFMSE total score at 12 months
Assess the efficacy of apitegromab compared with placebo based on the number of patients with clinical improvement in patients 2 through 21 years old	Proportion of patients with ≥3-point change from Baseline in the HFMSE total score at 12 months
Assess the efficacy of apitegromab compared with placebo by measuring changes in upper limb function between Baseline and the end of the Treatment Period using the RULM in patients 13 through 21 years old and in patients 2 through 21 years old	Change from Baseline in RULM total score at 12 months
Assess the efficacy of apitegromab compared with placebo by measuring changes in number of WHO motor development milestones in patients 13 through 21 years old and in patients 2 through 21 years old	Change from Baseline in number of WHO motor development milestones attained at 12 months
Further assess the efficacy of apitegromab compared with placebo by evaluating changes in additional motor function outcome measures and changes in HFMSE at other prespecified time points in patients 13 through 21 years old and in patients 2 through 21 years old	Proportion of patients achieving various magnitudes of change in HFMSE score from Baseline at 12 months • Proportion of patients achieving various magnitudes of change in RULM score from Baseline at 12 months • Proportion of patients who attain a new WHO motor development milestone relative to Baseline at 12 months • Change from Baseline in HFMSE total score at other prespecified time points

Objectives	Endpoints
	• Change from Baseline in RULM total score at other prespecified time points • Change from Baseline in number of WHO motor development milestones attained at other prespecified time points
Assess the efficacy of apitegromab compared with placebo by measuring changes from Baseline in motor function across the Treatment Period using the HFMSE in patients 13 through 21 years old and in patients 2 through 21 years old	Change from Baseline in HFMSE total score across time during the 12-month Treatment Period
Assess the time to stabilization of effect of apitegromab compared with placebo using the HFMSE in patients 13 through 21 years old	Time to decline (at least a -3-point change from Baseline in HFMSE total score) compared between apitegromab and placebo
Assess the time to therapeutic effect of apitegromab compared with placebo using the HFMSE in patients 2 through 21 years old	Time to therapeutic effect (≥1-point change from Baseline in HFMSE total score) compared between apitegromab and placebo
Evaluate the effects of apitegromab on fatigability, caregiver-reported disability, and suicidal ideation and behavior in patients 13 through 21 years old and in patients 2 through 21 years old	• Change from Baseline in PEDI-CAT • Change from Baseline in PROMIS Fatigue Questionnaire • Change from Baseline in ACEND • Change from Baseline in C-SSRS

Abbreviations: ACEND, Assessment of Caregiver Experience with Neuromuscular Disease; C-SSRS, Columbia-Suicide Severity Rating Scale; HFMSE, Hammersmith Functional Motor Scale - Expanded; PEDI-CAT, Pediatric Evaluation of Disability Inventory Computer Adaptive Test; PROMIS, Patient-Reported Outcomes Measurement Information System; RULM, Revised Upper Limb Module; WHO, World Health Organization.

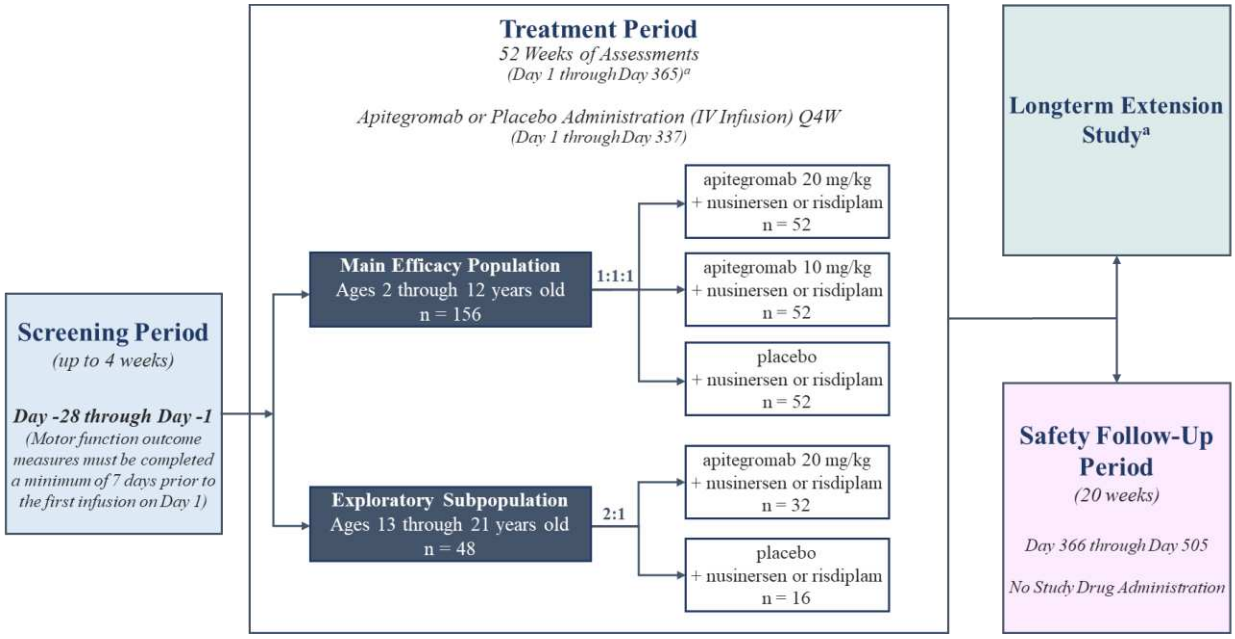
3. STUDY DESIGN

3.1. Brief Description

This randomized, double-blind, placebo-controlled Phase 3 trial (SRK-015-003 [SAPPHIRE]) will be conducted at approximately 55 to 60 trial sites globally to evaluate the efficacy and safety of apitegromab compared with placebo as an adjunctive therapy to nusinersen or risdiplam in nonambulatory patients with later-onset SMA. Patients in the MEP will be randomized to receive apitegromab (10 mg/kg or 20 mg/kg) or matching placebo by intravenous (IV) infusion. Patients in the EXP will be randomized to receive apitegromab 20 mg/kg or matching placebo by IV infusion. As shown in Figure 1, the trial will include Screening, Treatment, and Safety Follow-up Periods. The total trial participation time for a patient will consist of approximately 4 weeks for Screening, 52 weeks (ie, 12 months) of trial visits, and 20 weeks of safety follow-up for a total duration of approximately 76 weeks (approximately 18 months).

The schedule of activities/assessments (SoA) is provided in Section 14.1.

Figure 1: Overall Study Design



Abbreviations: IV, intravenous; Q4W, every 4 weeks.

[a] Patients, as well as the Investigator and site personnel, will remain blinded to the treatment assignment until the completion of the extension trial to minimize the bias in measures assessed in the extension trial.

Approximately 204 male and female patients with later-onset SMA will be randomized into either the MEP (2 through 12 years old at Screening) or the EXP (13 through 21 years old at Screening). Patients in the MEP will be randomized separately from the patients in the EXP.

Patients, as well as the investigator and site personnel, will remain blinded to the treatment assignment until the completion of the open label extension trial to minimize the bias in measures assessed in the open label extension trial.

Patients will be monitored throughout the trial for safety. Data will be reviewed on an ongoing basis by the Medical Monitor, the IDMC, and the Sponsor.

3.2. Determination of Sample Size

A sample size of 50 patients each in the apitegromab 20 mg/kg group and the placebo group from MEP would yield at least 80% power to detect a mean (± standard deviation) difference of 3 ± 5 points between the apitegromab 20 mg/kg group and the placebo group in the change from Baseline in the HFMSE total score, at a two-sided alpha (α) level of 0.05 and assuming a drop off rate of 5% after randomization.

With 50 patients in the apitegromab 10 mg/kg group, a sample size of 100 patients in the apitegromab combined dose (10 mg/kg and 20 mg/kg combined) group and 50 patients in the placebo group would yield at least 90% power to detect a mean (± standard deviation) difference of 3 ± 5 points between the apitegromab combined dose group and the placebo group.

With the Hochberg procedure (Hochberg 1988) to control for multiplicity of the 2 hypotheses of the superiority of apitegromab 20 mg/kg to placebo and the superiority of apitegromab combined dose (10 mg/kg and 20 mg/kg combined) to placebo specified in Section 6.1 as the primary confirmatory test, the power to reject at least 1 of the hypotheses is approximately 90%.

For the MEP, approximately156 patients will be randomized such that approximately 150 evaluable patients complete the Visit 14 assessments.

For the EXP, a maximum of 48 patients will be randomized. This sample size is based on practical considerations.

3.3. Treatment Assignment and Blinding

Patients in the MEP and the EXP will be randomized separately using a centralized Interactive Web-based Randomization System (IWRS) per Study Operations Manual and the IWRS Quick Reference Guide.

- For the MEP, patients will be randomized 1:1:1 double-blind to receive apitegromab 10 mg/kg, apitegromab 20 mg/kg, or placebo every 4 weeks during the 52-week (ie, 12-month) Treatment Period. Randomization for the MEP will be stratified by type of SMN therapy (ie, nusinersen or risdiplam) and age at initiation of SMN therapy (≥5 and <5 years).
- For the EXP, patients will be randomized 2:1 double-blind to receive apitegromab 20 mg/kg or placebo every 4 weeks during the 52-week (ie, 12-month) Treatment Period. Randomization for the EXP will be stratified by type of SMN therapy (ie, nusinersen or risdiplam).

The Sponsor, patients, caregivers, investigators, and site personnel will be blinded to treatment assignments throughout the study conduct. Access to the unblinded information is restricted and limited to only selected personnel based on their responsibilities and roles per Unblinded Team Management Plan. The unblinded personnel includes pharmacists, unblinded site staff, unblinded safety team, laboratories testing the PK/PD/ADA, unblinded clinical pharmacology team, unblinded Sponsor personnel, IDMC members, and vendor unblinded team. For all personnel needing to be unblinded, the list of personnel and the reason for unblinding will be documented along with the date of unblinding.

In the event of a drug-related, serious, unexpected adverse event (AE), designated unblinded Sponsor personnel may provide a patient's treatment assignment for the purpose of regulatory

authority agency reporting. In the event of a drug-related serious adverse event (SAE), the Investigator may, if deemed medically necessary to provide patient care, obtain the patient's treatment assignment from the IWRS system.

In case of an emergency, the Investigator will determine if unblinding of a patient's treatment assignment is warranted. If a patient's treatment assignment is to be unblinded, the Sponsor must be notified as soon as possible but no later than 24 hours after breaking the blind. The date and reason that the blind was broken must be recorded in the source documentation and electronic case report form (eCRF), as applicable.

For planned unblinding, such as IDMC meetings and final analysis, an authorization will be provided to release the treatment assignment for each planned unblinding.

4. ANALYSIS SETS

The definitions of each analysis set with the corresponding analysis treatment and patient population are provided in the Table 5.

Table 5: Definition of Analysis Sets

Analysis Set	Definition	Analysis Treatment	Patient Population
Randomized Set	All patients who were randomized	According to the treatment patients were randomized	Pooled Population
Intention-to-Treat (ITT) Set	All randomized patients in the MEP, even if the patient did not receive the correct treatment as assigned, or otherwise did not follow the protocol.	According to the treatment patients were randomized	MEP
Modified Intention-to-treat (MITT) Set	All ITT Set patients who received at least one dose of study drug and had at least one postbaseline evaluable HFMSE assessment.	According to the treatment patients were randomized	MEP
Per Protocol (PP) Set	All MITT Set patients who have no major protocol violation affecting the efficacy.	According to the treatment patients were randomized	MEP
Exploratory Subpopulation (EXP) Set	All randomized patients in the EXP and who received at least one dose of study drug.	According to the treatment patients were randomized	EXP
Full Analysis Set	All randomized patients who received at least 1 dose of study drug.	According to the treatment patients were randomized	Pooled Population
Safety Set	All randomized patients who receive at least 1 dose of study drug.	According to the actual treatment patients received	Pooled Population
PK Analysis Set	All randomized patients who receive at least 1 dose of study drug and had at least 1 quantifiable PK result.	According to the actual treatment patients received	Pooled Population
PD Analysis Set	All randomized patients who received at least 1 dose of study drug and had at least 1 evaluable circulating latent myostatin concentration	According to the actual treatment patients received	Pooled Population
Immunogenicity Analysis Set	All randomized patients who receive at least 1 dose of study drug and have at least 1 evaluable result	According to the actual treatment patients received	Pooled Population

Abbreviations: EXP, Exploratory Subpopulation; ITT, Intention-to-Treat; MEP, Main Efficacy Population; MITT, Modified Intention-to-Treat; PD, pharmacodynamic; PK, pharmacokinetic; PP, per protocol

For patients in the Intention-to-Treat (ITT) Set who were randomized but not treated or have been randomized and treated but had no postbaseline efficacy assessment, missing data will not be imputed, unless otherwise specified.

When all patients in the ITT Set received at least one dose of study drug and had at least one postbaseline evaluable HFMSE assessment, Modified Intention-to-Treat (MITT) Set is the same as ITT Set. The MITT Set is used as the main efficacy set when performing the primary efficacy analysis.

The study enrollment was completed when the current version of the SAP was developed, and all randomized patients received at least one dose of study drug (placebo or apitegromab at dose levels of 10 mg/mg or 20 mg/kg) and had at least one postbaseline evaluable HFMSE assessment. Therefore, the Randomized Set is the same as the Full Analysis Set, and the MITT Set is the same as ITT Set.

Criteria for exclusion from the Per-Protocol (PP) Set are listed as follows:

- Patients without evaluable HFMSE total score at 12-month
- Patients with major protocol deviation for Inclusion/Exclusion Criteria
- Patients with major protocol deviation of required primary assessment per protocol was not performed (HFMSE)
- Patients with major protocol deviation of "Unblinded personnel performs primary and key secondary assessments (HFMSE, RULM, or WHO MM)" at V14
- Patients with major protocol deviation of "Type or regimen of background therapy changed (replacing/adding/removing/adjustment of the therapy dose and/or frequency) during the trial"
- Patients with major protocol deviation of "Subject was not correctly stratified"
- Patients with major protocol deviation of "Patient was treated with apitegromab but not randomized"
- Patients with major protocol deviation of "Patient received an incorrect treatment (apitegromab instead of placebo or other way around)"
- Patients with major protocol deviation of "Patient received an incorrect treatment (incorrect dose of apitegromab per assigned arm)"

When all patients received the study drug same as they were randomized, the Full Analysis Set is the same as the Safety Set. The PK Analysis Set, the PD Analysis Set, and the Immunogenicity Analysis Set are subsets of the Safety Set.

5. GENERAL ANALYSIS

5.1. General Consideration

Summary statistics and statistical analyses will only be presented for those data detailed in this SAP and for the relevant analysis datasets specified in this SAP. Relevant data will be included in listings and presented by population (ie, MEP or EXP), by treatment, by patient and by visit where applicable.

The efficacy analyses will be conducted, including efficacy data collected during the Screening Period and Treatment Period. Efficacy measurements collected during the Safety Follow-up Period will be provided where applicable.

The safety analyses will be conducted, including all safety data collected. If there are any patients still being followed in the Safety Follow-up Period when all patients completed the Treatment Period and unblinding occurs, data collected after unblinding will be provided where applicable.

The statistical analyses will be tested at a two-sided 0.05 critical level, and 95% confidence interval (CI) will be presented where applicable.

Continuous variables will be summarized using the number of observations (n), mean, standard deviation (SD), median, minimum, and maximum. Categorical variables will be summarized using the number of observations (n), frequency, and percentage.

All analyses will be conducted using Statistical Analysis System (SAS®; Version 9.4 or later; Cary, North Carolina).

5.1.1. Analysis Groups Definition

In general, for analyses only including patients in the MEP, data will be presented by the 4 analysis groups defined below, unless otherwise specified:

- *Placebo*: This includes MEP patients who are in the placebo group in the relevant analysis sets.
- *10 mg/kg*: This includes MEP patients who are in the apitegromab 10 mg/kg group in the relevant analysis sets.
- *20 mg/kg*: This includes MEP patients who are in the apitegromab 20 mg/kg group in the relevant analysis sets.
- *SRK-015*: This includes MEP patients who are in the apitegromab 10 mg/kg or 20 mg/kg group in the relevant analysis sets.

In general, for analyses only including patients in the EXP, data will be presented by 2 analysis groups defined below, unless otherwise specified:

- *Placebo*: This includes EXP patients who are in the placebo group in the relevant analysis sets.
- *20 mg/kg*: This includes EXP patients who are in the apitegromab 20 mg/kg group in the relevant analysis sets.

In general, for analyses of Pooled Population that includes patients in both MEP and EXP, data will be presented by combining the MEP and EXP, and these analysis groups are defined below, unless otherwise specified:

- *Placebo*: This includes patients who are in the placebo group in the relevant analysis sets.
- *10 mg/kg*: This includes patients who are in the apitegromab 10 mg/kg group in the relevant analysis sets.
- *20 mg/kg*: This includes patients who are in the apitegromab 20 mg/kg group in the relevant analysis sets.
- *SRK-015*: This includes patients who are in the apitegromab 10 mg/kg or 20 mg/kg group in the relevant analysis sets.
- *Total*: This includes patients who are in all groups (ie, placebo or apitegromab 10 mg/kg or 20 mg/kg group) in the relevant analysis sets. The total column will be presented in the analyses of demographic, baseline characteristics, medication, and adverse event where applicable.

Table 6 summarizes the analysis groups for each patient population in general.

Table 6: Analysis Groups Presentation

Main Efficacy Population (MEP)	Exploratory Subpopulation (EXP)	Pooled Population
Placebo	Placebo	Placebo
10 mg/kg	20 mg/kg	10 mg/kg
20 mg/kg		20 mg/kg
SRK-015		SRK-015
		Total

5.1.2. Key Definitions

Study Day 1 is defined as the day when the first dose of study drug is taken, as recorded on the Study Drug Administration Case Report Form (CRF) page.

Study Days are calculated relative to Study Day 1. For an event that occurred after Study Day 1, the study day is calculated as (visit/event date minus Study Day1 plus 1). For an event that occurred before Study Day 1, the study day is calculated as (visit/event date minus Study Day1).

Baseline is generally defined as the last nonmissing measurement prior to the first infusion of study drug (ie, measurement has date/time that is either before or the same as the start date/time of the first infusion of study drug) except for the efficacy (ie, HFMSE, RULM, WHO, PEDI-CAT, PROMIS, and ACEND) and C-SSRS assessments. The Baseline for the efficacy assessments is defined as the last nonmissing measurement prior to or on the day of the first infusion of study drug (ie, measurement has date that is either before or the same as the start date of the first infusion of study drug). The Baseline for the C-SSRS is defined as the last nonmissing measurement for lifetime response from the Baseline/Screening form (ie, either the C-SSRS "Baseline/Screening" version used for patients ≥6 years old or the children's version of the C-SSRS "Baseline/Screening" used for patients who are 4 to 5 years old) prior to or on the

day of the first infusion of study drug (ie, measurement has date that is either before or the same as the start date of the first infusion of study drug).

5.1.3. General Data Imputation of Missing Items

5.1.3.1. HFMSE Total Score

In general, when calculating the HFMSE total score, if 20% or fewer (ie, ≤6) items are missing, then these items will be imputed to be 0 (unable) when summing all 33 items. If greater than 6 items are missing, then the total score will be set to be missing.

5.1.3.2. RULM Total Score

In general, when calculating the RULM total score, if 3 or fewer items are missing (ie, scores for both hands are missing for ≤3 items), then these items will be imputed to be 0 when summing all 19 scorable items. If greater than 3 items are missing, then the total score will be set to be missing.

5.1.3.3. Number of WHO Motor Development Milestones Attained

In general, when calculating the number of WHO motor development milestones attained, the milestone with a score of 1 will be counted.

To determine whether a new WHO motor development milestone is attained relative to baseline at a visit, the following steps will be applied.

- Step 1: A response of "3 - Yes" will be considered a score of 1 for that individual milestone, and a response of "1 - No (Inability)" will be considered a score of 0 for that individual milestone. If the response is neither "1 - No (Inability)" nor "3 - Yes", the individual milestone will be first set as missing.
- Step 2: For each individual milestone, the last nonmissing value prior to or on the day of the first infusion of study drug will be considered as baseline. If the individual milestone baseline is still missing, it will be imputed using the median value from patients with observed baseline values in the same patient population and randomization stratification. If the median value is 0.5, the missing individual milestone baseline will be imputed as 1. If all patients in the same patient population and randomization stratification have missing values, the missing individual milestone baseline will be imputed as 0 for all patients in that particular population and randomization stratification.
- Step 3: If an individual milestone is missing at a postbaseline analysis visit other than V14, the missing value will be imputed by using the worst value from the same patient at all other visits with observed values. If an individual milestone is missing at V14, it will be imputed using the lowest value from patients with observed values at V14 in the same patient population and randomization stratification.

5.1.4. Visit Windows

In general, data will be summarized or analyzed by visit for summary statistics and statistical analyses according to the scheduled visit as outlined in the protocol.

For primary and key secondary endpoints at postbaseline visits, when the scheduled visit assessment is not evaluable, data collected on the unscheduled visits or Early Termination (ET) Visit may be mapped to an appropriate analysis visit using the window scheme shown in Table 7. If there are 2 or more unscheduled assessments available in the same analysis window for a patient, the unscheduled assessment that is closest to the target visit day will be used for the analysis. If there are 2 or more unscheduled assessments in the same analysis window with the same distance from the target visit day, the earlier/earliest assessment will be used.

Table 7: Visit Windows for Primary and Key Secondary Endpoints

Scheduled Visit	Target Study Day	Analysis Visit Window in Study Days	
		Patients Who Entered the Extension Trial	Patients Who Do Not Enter the Extension Trial
Baseline (V1)	1	≤1	
V3	57	[2, 85]	
V5	113	[86, 141]	
V7	169	[142, 197]	
V9	225	[198, 253]	
V11	281	[254, 323]	
V14/EOS/EOT	365	≥324	[324, 379]
V15	393	N/A	[380, 407]
V16	421	N/A	[408, 463]
V17/EOS/ET	505	N/A	≥464

Abbreviations: EOS, end of study; EOT, end of treatment; ET, early termination; N/A, not applicable; V, visit.

5.2. Patient Disposition

The summary of patient disposition will be provided by treatment and overall using the Randomized Set. This summary will include the number of patients randomized. The summary will include the number and percentage for the following categories where the percentage is based on the number of randomized patients:

- Patients dosed.
- Treatment completion status and primary reason for treatment discontinuation.
- Treatment Period completion status and primary reason for Treatment Period discontinuation.
- Study completion status and primary reason for study discontinuation.

In addition, the number and percentage in each analysis set (defined in Section 4) will be summarized by treatment. If any patient received the study drug different from the assigned study drug, repeat the summary of patient disposition using the Safety Set.

The listing of patient disposition and inclusion of each analysis set at the patient level will be provided.

5.3. Demographic and Baseline Characteristics

Descriptive summaries of demographic and Baseline characteristics will be summarized by treatment using the Randomized Set. This summary will include age, height, weight, weight percentile, body mass index (BMI), BMI percentile, sex, race, ethnicity, type of SMN therapy at randomization (ie, nusinersen or risdiplam), age at initiation of SMN therapy (<5 or ≥5 years), and region (North America or Europe). Race will not be collected in France and Germany and therefore presented as not reported. Ethnicity will not be collected in France and therefore presented as not reported. The type of SMN therapy at randomization and the age at initiation of SMN therapy will be summarized according to the randomization stratum, with mis-stratification denoted in the footnote. Mis-stratified patients will be listed.

If any patient received study drug different from what was assigned (ie, Safety Set is different from Randomized Set), the summary table will also be repeated using the Safety Set. Age collected at Screening from the eCRF will be used. Height will be used in summary for all individuals who are able to independently stand. Surrogate height will be estimated using ulna length and used in summary for the individuals who are nonambulatory or need standing support. The Gauld's equation (Gauld 2004) will be used to derive the height (cm) if only ulna length is collected. The derivations for estimated height in centimeters (cm) for males and females are defined as follows:

Estimated height (Male) = (4.605 × ulna length) + [1.308 × min (age, 18)] + 28.003

Estimated height (Female) = (4.459 × ulna length) + [1.315 × min (age, 18)] + 31.485

Height will be summarized in centimeters (cm). Weight will be summarized in kilograms (kg).

BMI (kg/m^2) will be calculated as weight (kg)/[height (cm) / 100]^2.

The listings of subject demographic data and Baseline characteristics at the patient level will be provided.

5.4. Baseline Disease Characteristics

Descriptive summaries of Baseline SMA disease characteristics will be summarized by treatment and overall using the Randomized Set. The disease characteristics include age at diagnosis of SMA (years), time since diagnosis of SMA (years), age at SMA onset (years), age at initiation of SMN therapy (years), number of SMN2 gene copies, SMA type (Type 2 or Type 3), disease history of scoliosis status (yes or no), baseline contractures status (yes or no), severe contractures present at least one location, location of severe contractures (Hip Flexors, Knee Flexors, Ankle Planter Flexors, Elbow Flexors, or Forearm pronators), disease history of respiratory support status (yes or no), number of SMN therapies ever treated (1 or 2), duration on SMN therapy (ie, nusinersen or risdiplam) prior to study drug exposure, WHO milestone status when the 1st SMN therapy was started, Baseline HFMSE, Baseline WHO milestone, Baseline WHO milestone status and Baseline RULM. The derivations for time since diagnosis of SMA in years and duration on the SMN therapy prior to study drug exposure are defined as follows:

Time since diagnosis of SMA (years) = Age at screening – Age at diagnosis of SMA

Duration on the SMN therapy prior to study drug exposure (years) = (Date of the first dose of study drug – Date of the first dose of SMN therapy + 1) / 365.25

If there are any patients who received study drug different from what was assigned, the summary table will be repeated using the Safety Set.

A listing will be provided for the Baseline Disease Characteristics at the patient level.

5.5. Medical History

Medical history information will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) and summarized by treatment and overall for the Randomized Set. Summaries will be ordered alphabetically by system organ class (SOC) and then, within an SOC, alphabetically by preferred term (PT). A patient will be counted once for each SOC and counted once for each PT within the SOC. If there are any patients who received study drug different from what was assigned, the summary table will also be repeated using the Safety Set.

The listing of medical history data at the patient level will be provided.

5.6. Medication

5.6.1. Prior and Concomitant Medication

All medications will be collected from the date the informed consent form (ICF) is signed through the last trial visit.

Prior medications/therapies/devices are defined as any medications/therapies/devices (except for the SMN therapies of nusinersen and risdiplam, which will be summarized separately) stopped before the administration of the first dose of the study drug. Concomitant medications are defined as any medications (except for the SMN therapies of nusinersen and risdiplam) taken concurrently while on study drug (ie, medications that started prior to the end of study and either stopped after the administration of the first dose of study drug or ongoing).

For a record with a completed missing start/end date, the ongoing status and/or whether the medication/therapy/device is taken prior to signing consent will be applied to determine whether it is concomitant.

For a record with a partial start/end date, the year/month of the partial date will be compared to date of the first infusion to determine whether it is concomitant.

All medications will be coded using WHO Drug Global B3 dictionary.

Prior medications will be summarized using the Anatomical Therapeutic Chemical (ATC) Classification Level 2 and PT by treatment and overall for the Safety Set. If a patient has taken a medication more than once, the patient will be counted only once. A listing of prior medications at the patient level will be provided.

Concomitant medications (except for the SMN therapies of nusinersen and risdiplam) will be summarized using the ATC Classification Level 2 and PT by treatment and overall using the Safety Set. If a patient has taken a medication more than once, the patient will be counted only once. A listing of concomitant medications at the patient level will be provided.

Concomitant therapies or interventional procedures that are medically indicated for any AEs the patient has during the trial or that are provided as part of standard supportive care for the patient will be recorded if performed.

Concomitant therapies will only be provided in data listing.

5.6.2. Prior and Concomitant SMN Therapy

Prior SMN therapies (ie, nusinersen or risdiplam) will be summarized by treatment and overall using the Safety Set. This summary will include the number of SMN therapies treated (1 or 2) prior to the study enrollment, number of patients receiving nusinersen at Screening, and number of patients receiving risdiplam at Screening.

In addition, the primary reason for changing the SMN therapy prior to the study enrollment will be presented for those patients who have been exposed to 2 SMN therapies prior to the study enrollment.

The details of prior SMN therapy at the patient level will be listed.

In addition, the concomitant SMN therapy at the patient level will be listed, which includes the start date and end date/ongoing status with the dose level.

5.7. Study Drug Exposure

Study drug exposure including the number of infusions received, number of completed infusions, duration on study drug, and duration of exposure will be summarized as a continuous variable by treatment and overall using the Safety Set. The duration of exposure is defined as follows.

Duration of Exposure (weeks) = (Last dose date of study drug – First dose date of study drug + 1) / 7

5.8. Study Drug Compliance

Study drug compliance will be calculated as follows:

$$100 \times \frac{\text{Number of infusions received}}{\text{Number of infusions expected to be received}}$$

Study drug compliance will be considered as a continuous variable and summarized by treatment and overall for the Safety Set. The number and percentages of patients with study drug compliance categories (ie, <80% or ≥80%) will also be provided.

5.9. Protocol Deviations

If any issue relating to the safety of a study patient arises that requires a deviation from the protocol, the study unit through the Investigator may immediately make such a deviation. If there is a need for such a deviation, the study unit must notify the Sponsor and the responsible person in Institutional Review Board (IRB)/Independent Ethics Committee (IEC). The nature and reasons for the protocol deviations will be recorded. All protocol deviations will be captured by the study site personnel in the Rave clinical trial management system and categorized as major or minor.

All major protocol deviations, including those being used as criteria for exclusion from the PP Set, observed during the conduct of the study will be listed. Major protocol deviations will be summarized by treatment and overall for each protocol deviation category for the Randomized Set and the MITT Set.

6. EFFICACY ANALYSIS

6.1. Primary Efficacy Endpoint

The primary objective of the study is to assess the efficacy of apitegromab compared with placebo using the HFMSE in patients 2 through 12 years old.

The alternative hypotheses of the primary objective are as follows:

- Apitegromab 20 mg/kg is superior to placebo for change from Baseline in HFMSE total score at 12 months, and
- Apitegromab combined dose (10 mg/kg and 20 mg/kg combined) is superior to placebo for change from Baseline in HFMSE total score at 12 months.

6.1.1. Main Estimand

The primary efficacy estimand (ICH E9(R1) Addendum 2019) for the primary efficacy endpoint "Change from Baseline in HFMSE total score at 12 months" consists of two components: (1) the main estimand for the comparison between apitegromab 20 mg/kg and placebo and (2) the main estimand for the comparison between apitegromab combined dose (10 mg/kg and 20 mg/kg combined) and placebo, as defined below:

- Treatment: Apitegromab or placebo administered every 4 weeks by IV infusion during the 52-week (ie, 12-month) Treatment Period.
  - Apitegromab 20 mg/kg for comparison between apitegromab 20 mg/kg and placebo
  - Apitegromab 10 mg/kg or 20 mg/kg for comparison between apitegromab combined dose and placebo
- Population: Patients in the MITT Set.
- Variable: Change from Baseline in HFMSE total score at 12 months.
- Population-level Summary: The least squares (LS) mean difference in change from Baseline.
- Intercurrent events (ICEs) and Strategies for Handling ICEs*:*
  - Premature stopping of study treatment due to AE: Treatment policy strategy will be applied where values after the ICEs will be used.
  - Premature stopping of the study treatment due to death: Composite variable strategy will be applied where values after the ICEs will be imputed (see Section 6.1.4 for further details in the event death occurs).
  - Premature stopping of study treatment due to reasons other than AE and death: Treatment policy strategy will be applied where values after the ICEs will be used.
  - Missed 3 or more consecutive doses impacted by coronavirus disease 2019 (COVID-19): Treatment policy strategy will be applied where values after the ICEs will be used.

- Missed 3 or more consecutive doses other than impact by COVID-19: Treatment policy strategy will be applied where values after the ICEs will be used.
- Had scoliosis/spinal surgery: Treatment policy strategy will be applied where values after the ICEs will be used. (Young 2020).
- Stopping receiving SMN therapy for longer than 6 months of nusinersen or 10 days of risdiplam without switching to another SMN therapy: Treatment policy strategy will be applied where values after the ICEs will be used.
- SMN therapy switched during the study: Treatment policy strategy will be applied where values after the ICEs will be used.

If a patient had more than 1 ICE, the first occurrence will be considered.

Table 8 also summarizes the definitions of the main estimands for the primary endpoint and the key secondary endpoints.

6.1.2. Supportive Estimand

The supportive estimand for the comparison between apitegromab 20 mg/kg and placebo and the supportive estimand for the comparison between apitegromab combined dose (10 mg/kg and 20 mg/kg combined) and placebo are similar to the main estimands defined in Section 6.1.1 but by applying the following strategies to handle the ICEs.

ICEs and Strategies for Handling ICEs:

- Premature stopping of study treatment due to AE: Composite variable strategy will be applied where values after the ICEs will be imputed.
- Premature stopping of the study treatment due to death: Composite variable strategy will be applied where values after the ICEs will be imputed.
- Premature stopping of study treatment due to reasons other than AE and death: Treatment policy strategy will be applied where values after the ICEs will be used.
- Missed 3 or more consecutive doses impacted by COVID-19: Hypothetical strategy will be applied where values after the ICEs will be set as missing.
- Missed 3 or more consecutive doses other than impact by COVID-19: Treatment policy strategy will be applied where values after the ICEs will be used.
- Had scoliosis/spinal surgery: Hypothetical strategy will be applied where values after the ICEs will be set as missing (Young 2020).
- Stopping receiving SMN therapy for longer than 6 months of nusinersen or 10 days of risdiplam without switching to another SMN therapy: Treatment policy strategy will be applied where values after the ICEs will be used.
- SMN therapy switched during the study: Treatment policy strategy will be applied where values after the ICEs will be used.

6.1.3. Data Imputation

Missing values (ie, HFMSE total scores) will be imputed by different imputation approaches based on the reasons of missing and/or different assumption of missing mechanisms.

The missing values after ICEs of "Premature stopping of the study treatment due to death" will be imputed using probabilistic sampling from the lowest 10% values within in each treatment group per visit for both the main and supportive estimands before conducting the primary analysis and the sensitivity analysis.

The missing values after ICEs of "Premature stopping of study treatment due to AE" will be imputed using probabilistic sampling from the lowest 10% values within in each treatment group per visit for supportive estimand before conducting the primary analysis and the sensitivity analysis.

The missing values after ICEs of "Had scoliosis/spinal surgery" will be handled by mixed effects model with repeated measurement (MMRM) for the primary analysis under the missing-at-random (MAR) assumption. For the sensitivity analysis, the missing values will be imputed using probabilistic sampling from the lowest 10% values within in each treatment group per visit.

With the exception of missing values after ICEs of "Premature stopping of the study treatment due to death", "Premature stopping of study treatment due to AE" and "Had scoliosis/spinal surgery" for the corresponding estimand, other missing values (regardless of setting as missing or were missing) after ICEs will either be handled by MMRM for the primary analysis under the MAR assumption or be imputed by multiple imputation (MI) assuming missing-not-at-random (MNAR) for the sensitivity analysis. More specifically, the Pattern-Mixture Model (PMM) approach will be implemented to create a control-based pattern imputation (Little 1993, Ratitch and O'Kelly 2011). The sample SAS code applying the MI under the MNAR assumption is provided in Section 14.2.

For intermediate missing values, they will be imputed via a Markov chain Monte Carlo (MCMC) method before applying the MI methods imputing the monotone missing. The MCMC method is applied under the assumption of multivariate normal distribution, which will produce a monotone missing pattern [data with only terminal missing and no intermediate missing] (Li 1988, Schafer 1997). The sample SAS® code applying MCMC is provided in Section 14.2 before applying the MI under the MNAR assumption.

Any other missing values will either be handled by MMRM for the primary analysis under the MAR assumption or be imputed by multiple imputation (MI) assuming MAR for the sensitivity analysis.

If any death occurred and when randomization based method for covariance and stratified adjustment of win odds (accounts for ties) for ordinal outcome (Weideman 2022, Kowalewski 2023) is conducted, the missing values will be imputed by MI assuming MAR. The sample SAS code applying the MI under the MAR assumption is provided in Section 14.3.

6.1.4. Primary Analysis

The primary analysis for the main and supportive estimands will be performed using a restricted maximum likelihood model (REML) based MMRM model to analyze the change from Baseline in HFMSE total score at 12 months. The model will include the fixed effects of treatment, visit,

treatment-by-visit interaction, baseline HFMSE total score, baseline HFMSE total score-by-time interaction, type of SMN therapy at randomization (ie, nusinersen or risdiplam), and age at initiation of SMN therapy (≥5 and <5 years). An unstructured covariance structure will be used to model the within-patient variability. If the model with the unstructured covariance structure fails to converge, the heterogeneous first-order autoregression, AR(1), covariance structure will be used. The Kenward-Roger method will be used to estimate the degrees of freedom. The LS mean difference between apitegromab and placebo treatments groups at 12 months will be presented with the corresponding 95% CI and the p-value. The sample SAS code performing the MMRM is provided in Section 14.4.

If any death occurred, the p-value from the randomization based method for covariance and stratified adjustment of win odds (accounts for ties) for ordinal outcome will be presented as the primary analysis. The type of SMN therapy at randomization (ie, nusinersen or risdiplam) and the age at initiation of SMN therapy (≥5 and <5 years) will be included as the stratification variable of 4 strata (ie, combination of the type of SMN therapy at randomization and the age at initiation of SMN therapy). If all the values in one stratum within any of the analysis groups are the same, stratification factors will be reduced to 3 strata by combining the two strata with the least number of patients. If same value is still observed in any of the 3 strata, instead of further combining the strata, the type of SMN therapy at randomization (ie, nusinersen or risdiplam) will be used as the stratification factor. If same value is still observed in any strata of the type of SMN therapy at randomization (ie, nusinersen or risdiplam), rank based ANCOVA (Conover 1982) will be considered as alternative approach. The Baseline HFMSE total score will be included as baseline variable. The sample SAS code performing the randomization based method for covariance and stratified adjustment of win odds for ordinal outcome is provided in Section 14.5.

The descriptions of ICEs and strategies for handling ICEs for the primary endpoint and the continuous key secondary endpoints, as well as missing data handling in the primary analysis (ie, MMRM) and the sensitivity analysis using analysis of covariance (ANCOVA), are summarized in Table 9.

6.1.5. Sensitivity Analysis

A sensitivity analysis will be conducted for the main and supportive estimands by an ANCOVA model adjusting for treatment, type of SMN therapy at randomization (ie, nusinersen or risdiplam), age at initiation of SMN therapy (≥5 and <5 years), and Baseline HFMSE total score. The sensitivity analysis will be performed using the imputed datasets for each estimand, and the results will be combined based on Rubin's rules (Rubin 1987) assuming the estimated statistics from each imputed dataset follow a normal distribution. The sample SAS code performing the ANCOVA is provided in Section 14.6.

If any death occurred, the p-value applying the MMRM will be provided as additional sensitivity analysis.

If more than 5% of the MEP patients (>7 MEP patients) had ICEs, an additional sensitivity analysis will be conducted by applying a tipping-point analysis to evaluate the robustness of the estimates for the main estimand. The tipping-point analysis will be conducted using the MMRM model by adding a shift variable from a prespecified range to the imputed values in the apitegromab 10 mg/kg or 20 mg/kg groups assuming MNAR while the missing values in the

placebo group are assumed MAR. The details of performing the tipping-point analysis is provided in Section 14.7.

6.1.6. Supplementary Analysis

A supplementary analysis will be performed by repeating the primary analysis and sensitivity analysis for the main estimand using the PP Set.

A summary table of ICE(s) will be provided by treatment using the MITT Set and the PP Set.

If the primary analysis for the main and supportive estimands leads to different conclusions, an additional supplementary analysis will be conducted by applying the randomization-based method for covariance and stratified adjustment of win odds (accounts for ties) for ordinal outcome.

6.1.7. Key Secondary Efficacy Estimands and Analyses for Apitegromab 10 mg/kg

The main estimand for comparisons between apitegromab 10 mg/kg and placebo is considered as one of the key secondary efficacy estimands, and the primary analysis of the main estimand is one of the confirmatory tests in the hierarchical testing procedure (Section 6.3).

The estimands for the comparisons between apitegromab 10 mg/kg and placebo are similar to the estimands for the comparison between apitegromab 20 mg/kg and placebo. The differences in defining the estimands are listed below while other elements are the same as the estimands for the comparison between apitegromab 20 mg/kg and placebo defined in Section 6.1.1 and Section 6.1.2.

- Treatment: Apitegromab 10 mg/kg or placebo applied Q4W by IV infusion during a 52-week (ie, 12-month) Treatment Period.

The primary analyses, sensitivity analyses, and supplementary analyses for the comparison between apitegromab 10 mg/kg and placebo are the same as the primary, sensitivity and supplementary analyses for the comparison between apitegromab and placebo in Section 6.1.4, Section 6.1.5, and Section 6.1.6.

6.2. Key Secondary Efficacy Endpoints

6.2.1. Key Secondary Efficacy Endpoint: Change from Baseline in RULM Total Score at 12 Months

6.2.1.1. Main Estimand

The main estimand for the key secondary efficacy endpoint “Change from Baseline in RULM total score at 12 months” is similar to the main estimand of the primary endpoint defined in Section 6.1.1 using the variable of change from Baseline in RULM total score at 12 months to define the main estimand. The definition is also presented in Table 8 with that of the other main estimands. For those patients who did not complete the Baseline RULM assessment because of age restriction (ie, <30 months at the time of the Baseline assessment), no imputation will be applied for those subjects and these subjects will not be included in the analysis.

6.2.1.2. Supportive Estimand

The supportive estimand for this continuous key secondary endpoint is defined similarly to the supportive estimand of the primary endpoint defined in Section 6.1.2 with the different variable of change from Baseline in RULM total score at 12 months.

6.2.1.3. Data Imputation

The missing value imputation for this key secondary endpoint will be the same as the data imputation approaches applied on the primary endpoint described in Section 6.1.3. The missing RULM total scores will be imputed by different imputation approaches based on the reasons of missing and/or different assumption of missing mechanisms.

6.2.1.4. Primary Analysis

The primary analysis for the main and supportive estimands will be performed using the same MMRM model for the primary endpoint described in Section 6.1.4 to analyze the change from Baseline in RULM total score at 12 months.

6.2.1.5. Sensitivity Analysis

A sensitivity analysis will be performed for both estimands similar to the sensitivity analysis of the primary efficacy estimands described in Section 6.1.5.

6.2.1.6. Supplementary Analysis

A supplementary analysis will be performed by repeating the primary analysis and sensitivity analysis for the main estimand on the PP Set.

6.2.1.7. Analyses for Apitegromab 10 mg/kg

The main and supportive estimands for the comparison between apitegromab 10 mg/kg and placebo are similar to the comparison between apitegromab 20 mg/kg and placebo defined in Section 6.2.1.1 and Section 6.2.1.2 with the difference in the treatment of comparison.

The primary analyses, sensitivity analyses, and supplementary analyses for the comparison between apitegromab 10 mg/kg and placebo are the same as the primary, sensitivity, and supplementary analyses for the comparison between apitegromab 20 mg/kg and placebo with details in Section 6.2.1.4, Section 6.2.1.5, and Section 6.2.1.6.

6.2.1.8. Analyses for Apitegromab

The definitions of the main and supportive estimands for the comparisons between apitegromab combined dose (10 mg/kg and 20 mg/kg combined) and placebo are similar to the estimands for the comparison between apitegromab 20 mg/kg and placebo.

The primary analyses, sensitivity analyses, and supplementary analyses for the comparison between apitegromab and placebo are the same as the primary, sensitivity, and supplementary analyses for the comparison between apitegromab 20 mg/kg and placebo.

6.2.2. Key Secondary Efficacy Endpoint: Proportion of Patients With ≥3-Point Change From Baseline in HFMSE Total Score at 12 Months

6.2.2.1. Main Estimand

The main estimand for the key secondary efficacy endpoint of "Proportion of patients with ≥3-point change from Baseline in the HFMSE total score at 12 months" for the comparison between apitegromab 20 mg/kg and placebo is defined as follows:

- Treatment: Apitegromab 20 mg/kg or placebo applied Q4W by IV infusion during the 52-week (ie, 12-month) Treatment Period.
- Population: Patients in the MITT Set.
- Variable: Binary response variable (ie, 1 versus 0) with 1 defined as a patient with ≥3-point change from Baseline in the HFMSE total score at 12 months.
- Population-level summary: Odds ratio of patients with ≥3-point change from Baseline in HFMSE total score at 12 months.
- ICEs and strategies for handling ICEs:
  - Premature stopping of study treatment due to AE: Treatment policy strategy will be applied where values after the ICEs will be used.
  - Premature stopping of the study treatment due to death: Composite variable strategy will be applied where values after the ICEs will be imputed.
  - Premature stopping of study treatment due to reasons other than AE and death: Treatment policy strategy will be applied where values after the ICEs will be used.
  - Missed 3 or more consecutive doses impacted by COVID-19: Treatment policy strategy will be applied where values after the ICEs will be used.
  - Missed 3 or more consecutive doses other than impact by COVID-19: Treatment policy strategy will be applied where values after the ICEs will be used.
  - Had scoliosis/spinal surgery: Treatment policy strategy will be applied where values after the ICEs will be used (Young 2020).
  - Stopping receiving SMN therapy for longer than 6 months of nusinersen or 10 days of risdiplam without switching to another SMN therapy: Treatment policy strategy will be applied where values after the ICEs will be used.
  - SMN therapy switched during the study: Treatment policy strategy will be applied where values after the ICEs will be used.

Table 8 also summarizes the definitions of the main estimands for the primary endpoint and the key secondary endpoints.

6.2.2.2. Supportive Estimand

The supportive estimand for the key secondary efficacy endpoint of "Proportion of patients with ≥3-point change from Baseline in the HFMSE total score at 12 months" is defined similarly to the main estimand with the difference in the strategies specified to handle the ICEs. The

strategies to handle the ICEs for the supportive estimand for the key secondary efficacy endpoint are the same as those for the supportive estimand for the primary endpoint that are specified in Section 6.1.2.

6.2.2.3. Data Imputation

The missing binary response variable (ie, 1 versus 0) will be imputed by different imputation approaches based on the reasons of missing and/or different assumption of missing mechanisms.

The missing binary response variable (ie, 1 versus 0) after any ICEs of "Premature stopping of study treatment due to AE", "Premature stopping of the study treatment due to death", "Missed 3 or more consecutive doses other than impact by COVID-19", "Had scoliosis/spinal surgery" or "Stopping receiving SMN therapy for longer than 6 months of nusinersen or 10 days of risdiplam without switching to another SMN therapy" will be considered as 0 (defined as a patient with <3-point change from Baseline in the HFMSE total score at 12 months).

The missing binary response variable (ie, 1 versus 0) after any ICEs of either "Premature stopping of study treatment due to reasons other than AE and death", "Missed 3 or more consecutive doses impacted by COVID-19", or "SMN therapy switched during the trial" as well as any other missing values will be derived based on the imputed change from baseline in HFMSE total scores. The imputed change from baseline in HFMSE total scores will be imputed by MI under the MAR assumption. The sample SAS code applying the MI under the MAR assumption is provided in Section 14.3. For intermediate missing values, they will be imputed via an MCMC method before applying the MI methods imputing the monotone missing. The MCMC method is applied under the assumption of multivariate normal distribution, which will produce a monotone missing pattern [data with only terminal missing and no intermediate missing] (Li 1988; Schafer 1997).

The descriptions of ICEs and strategies for handling ICEs for the binary key secondary endpoint, as well as missing data handling, are summarized in Table 10.

6.2.2.4. Primary Analysis

The primary analysis for the main and supportive estimands will be performed using the logistic regression model, which include covariates of Baseline HFMSE total score, type of SMN therapy at randomization (ie, nusinersen or risdiplam), and age at initiation of SMN therapy (≥5 and <5 years). The odds ratio estimates will be presented with the corresponding 95% CIs and the p-values. The sample SAS code performing the logistic regression and combining the results from imputed datasets is provided in Section 14.8 (Ratitch 2013).

6.2.2.5. Sensitivity Analysis

The sensitivity analysis for the main and supportive estimands will be analyzed using the Cochran-Mantel-Haenszel (CMH) test (Mantel 1963) with stratification factors of type of SMN therapy at randomization (ie, nusinersen or risdiplam) and age at initiation of SMN therapy (≥5 and <5 years). The odds ratio estimates will be presented with the corresponding 95% CIs and the p-values. The sample SAS® code performing the CMH test and combining the results from imputed datasets is provided in Section 14.9 (Ratitch 2013).

6.2.2.6. Supplementary Analysis

A supplementary analysis will be performed by repeating the primary analysis and sensitivity analysis for the main estimand using the PP Set.

A frequency table of the number and percentage of patients with ≥3-point change from Baseline in HFMSE total score at each postbaseline visit (including the 12-month visit) will be provided by treatment using the MITT Set and the PP Set.

6.2.2.7. Analyses for Apitegromab 10 mg/kg

The definitions of the main and supportive estimands for the comparison between apitegromab 10 mg/kg and placebo are similar to the comparison between apitegromab 20 mg/kg and placebo defined in Section 6.2.2.1 and Section 6.2.2.2 with the difference in the treatment of comparison.

The primary analyses, sensitivity analyses, and supplementary analyses for the comparison between apitegromab 10 mg/kg and placebo are the same as the primary, sensitivity analyses, and supplementary analyses for the comparison between apitegromab 20 mg/kg and placebo in Section 6.2.2.4 and Section 6.2.2.5.

6.2.2.8. Analyses for Apitegromab

The definitions of the main and supportive estimands for the comparison between apitegromab combined dose (10 mg/kg and 20 mg/kg combined) and placebo are similar to the comparison between apitegromab 20 mg/kg and placebo defined in Section 6.2.2.1 and Section 6.2.2.2 with the difference in the treatment of comparison.

The primary analyses, sensitivity analyses, and supplementary analyses for the comparison between apitegromab and placebo are the same as the primary, sensitivity analyses, and supplementary analyses for the comparison between apitegromab 20 mg/kg and placebo in Section 6.2.2.4 and Section 6.2.2.5.

6.2.3. Key Secondary Efficacy Endpoint: Change from Baseline in Number of WHO Motor Development Milestones Attained at 12 Months

6.2.3.1. Main Estimand

The main estimand for the key secondary efficacy endpoint “Change from Baseline in number of WHO motor development milestones attained at 12 months” is similar to the main estimand of the primary endpoint defined in Section 6.1.1 but using the variable of change from Baseline in number of WHO motor development milestones attained at 12 months to define the main estimand. The definition is also presented in Table 8 with that of the other main estimands.

6.2.3.2. Supportive Estimand

The supportive estimand for this continuous key secondary endpoint is defined similarly to the supportive estimand of the primary endpoint defined in Section 6.1.2 with the different variable of change from Baseline in number of WHO motor development milestones attained at 12 months.

6.2.3.3. Data Imputation

The missing value imputation for this key secondary endpoint will be the same as the data imputation approaches applied on the primary endpoint described in Section 6.1.3. The missing number of WHO motor development milestones attained will be imputed by different imputation approaches based on the reasons of missing and/or different assumption of missing mechanisms.

6.2.3.4. Primary Analysis

The primary analysis for the main and supportive estimands will be performed using the same MMRM model for the primary endpoint described in Section 6.1.4 to analyze the change from Baseline in number of WHO motor development milestones attained at 12 months.

6.2.3.5. Sensitivity Analysis

A sensitivity analysis will be performed for both estimands similar to the sensitivity analysis of the primary efficacy estimands described in Section 6.1.5.

6.2.3.6. Supplementary Analysis

A supplementary analysis will be performed by repeating the primary analysis and sensitivity analysis for the main estimand on the PP Set.

6.2.3.7. Analyses for Apitegromab 10 mg/kg

The main and supportive estimands for the comparison between apitegromab 10 mg/kg and placebo are similar to the comparison between apitegromab 20 mg/kg and placebo defined in Section 6.2.3.1 and Section 6.2.3.2 with the difference in the treatment of comparison.

The primary analyses, sensitivity analyses, and supplementary analyses for the comparison between apitegromab 10 mg/kg and placebo are the same as the primary, sensitivity, and supplementary analyses for the comparison between apitegromab and placebo with details in Section 6.2.3.4, Section 6.2.3.5, and Section 6.2.3.6.

6.2.3.8. Analyses for Apitegromab

The definitions of the main and supportive estimands for the comparisons between apitegromab combined dose (10 mg/kg and 20 mg/kg combined) and placebo are similar to the estimands for the comparison between apitegromab 20 mg/kg and placebo. The primary analyses, sensitivity analyses, and supplementary analyses for the comparison between apitegromab and placebo are also the same as the primary, sensitivity, and supplementary analyses for the comparison between apitegromab 20 mg/kg and placebo.

Table 8: Definition of Main Estimands for Primary Endpoint and Key Secondary Endpoints

Estimand component	HFMSE	HFMSE	RULM[a]	WHO
	Continuous variable	Binary variable	Continuous variable	Continuous variable
Treatment	Apitegromab[b] or placebo administered every 4 weeks by IV infusion during a 52-week (ie, 12-month) Treatment Period			
Population	All patients in the MITT Set			
Variable	Change from Baseline in HFMSE total score at 12 months	Binary response variable (ie, 1 vs 0) with 1 defined as a patient with ≥3-point change from Baseline in HFMSE total score at 12 months	Change from Baseline in RULM total score at 12 months	Change from Baseline in number of WHO motor development milestones attained at 12 months
Population-level summary	LS mean difference in change from Baseline	Odds ratio of patients with ≥3-point change from Baseline in the HFMSE total score at 12 months	LS mean difference in change from Baseline	
Intercurrent events (ICEs)	- Premature stopping of study treatment due to AE. - Premature stopping of the study treatment due to death. - Premature stopping of study treatment reasons other than AE and death. - Missed 3 or more consecutive doses impacted by COVID-19. - Missed 3 or more consecutive doses other than impact by COVID-19. - Had scoliosis/spinal surgery. - Stopping receiving SMN therapy for longer than 6 months of nusinersen or 10 days of risdiplam without switching to another SMN therapy. - SMN therapy switch during the trial.			

Abbreviations: AE, adverse event; COVID-19, coronavirus disease 2019; HFMSE, Hammersmith Functional Motor Scale - Expanded; ICE, intercurrent event; IV, intravenous; LS, least squares; MITT, modified Intention-to-Treat; RULM, Revised Upper Limb Module; SMN, survival motor neuron; WHO, World Health Organization.

[a] For those patients who did not complete the Baseline RULM assessment because of age restriction (ie, <30 months at the time of the Baseline assessment), no imputation will be applied, and those patients will not be included in the analysis.

[b] Apitegromab 20 mg/kg for comparison between apitegromab 20 mg/kg and placebo; apitegromab 10 mg/kg or 20 mg/kg for comparison between apitegromab combined dose and placebo; apitegromab 10 mg/kg for comparison between apitegromab 10 mg/kg and placebo.

Table 9: Intercurrent Events, Strategies for Handling ICEs for the Primary Endpoint and the Continuous Key Secondary Endpoints as well as Missing Data Handling in the Primary Analysis and the Sensitivity Analysis

<table>
<tr><th rowspan="3">ICE</th><th colspan="2">Estimands: Proposed Strategy Handing the ICE</th><th colspan="2">Missing Data Handling</th></tr>
<tr><th rowspan="2">Main Estimand</th><th rowspan="2">Supportive Estimand</th><th colspan="2">Primary Endpoint and Continuous Key Secondary Endpoints</th></tr>
<tr><th>Primary Analysis (MMRM)</th><th>Sensitivity Analysis (ANCOVA)</th></tr>
<tr><td>Premature stopping of the study treatment due to death</td><td colspan="2">Composite variable strategy: Values after the ICEs will be imputed using probabilistic sampling from the lowest 10% values within in each treatment group per visit.</td><td colspan="2">The missing values will be imputed using probabilistic sampling from the lowest 10% values within in each treatment group per visit.</td></tr>
<tr><td>Had scoliosis/spinal surgery</td><td>Treatment policy strategy: values after the ICEs will be used.</td><td>Hypothetical strategy: Values after the ICEs will be set as missing.</td><td>The missing values will be handled via MMRM assuming MAR.</td><td>The missing values will be imputed using probabilistic sampling from the lowest 10% values within in each treatment group per visit.</td></tr>
<tr><td>Premature stopping of study treatment due to AE</td><td>Treatment policy strategy: Values after the ICEs will be used.</td><td>Composite variable strategy: Values after the ICEs will be imputed using probabilistic sampling from the lowest 10% values within in each treatment group per visit.</td><td rowspan="6">The missing values will be handled via MMRM assuming MAR.</td><td rowspan="6">The missing values will be imputed by MI assuming MNAR using the PMM.</td></tr>
<tr><td>Missed 3 or more consecutive doses impacted by COVID-19</td><td>Treatment policy strategy: Values after the ICEs will be used.</td><td>Hypothetical strategy: values after the ICEs will be set as missing.</td></tr>
<tr><td>Premature stopping of study treatment due to reasons other than AE and death</td><td colspan="2" rowspan="3">Treatment policy strategy: Values after the ICEs will be used.</td></tr>
<tr><td>Missed 3 or more consecutive doses other than impact by COVID-19</td></tr>
<tr><td>Stopping receiving SMN therapy for longer than 6 months of nusinersen or 10 days of risdiplam without switching to another SMN therapy</td></tr>
<tr><td>SMN therapy switched during the trial</td><td colspan="2">Treatment policy strategy: Values after the ICEs will be used.</td></tr>
</table>

Abbreviations: AE, adverse event; ANCOVA, analysis of covariance; COVID-19, coronavirus disease 2019; ICE, intercurrent event; MAR, missing-at-random; MI, multiple imputation; MMRM, mixed model repeated measures; MNAR, missing-not-at-random; PMM, pattern-mixture model; SMN, survival motor neuron.

Table 10: Intercurrent Events (ICEs), Strategies for Handling ICEs for the Binary Key Secondary Endpoint as well as Missing Data Handling

ICE	Estimands: Proposed Strategy Handing the ICE		Missing Data Handling
	Main Estimand	Supportive Estimand	
Premature stopping of the study treatment due to death	Composite variable strategy: Values after the ICEs will be imputed as 0.		The missing values will be imputed as 0.
Had scoliosis/spinal surgery	Treatment policy strategy: Values after the ICEs will be used.	Hypothetical strategy: Values after the ICEs will be set as missing.	
Premature stopping of study treatment due to AE	Treatment policy strategy: Values after the ICEs will be used.	Composite variable strategy: Values after the ICEs will be imputed as 0.	
Missed 3 or more consecutive doses impacted by COVID-19	Treatment policy strategy: Values after the ICEs will be used.	Hypothetical strategy: Values after the ICEs will be set as missing.	The missing binary response will be derived based on the imputed change from baseline in HFMSE total scores imputed by MI assuming MAR.
Missed 3 or more consecutive doses other than impact by COVID-19	Treatment policy strategy: Values after the ICEs will be used.		The missing values will be imputed as 0.
Stopping receiving SMN therapy for longer than 6 months of nusinersen or 10 days of risdiplam without switching to another SMN therapy			
Premature stopping of study treatment due to reasons other than AE and death			The missing binary response will be derived based on the imputed change from baseline in HFMSE total scores imputed by MI assuming MAR.
SMN therapy switched during the trial			

Abbreviations: AE, adverse event; COVID-19, coronavirus disease 2019; HFMSE, Hammersmith Functional Motor Scale - Expanded; ICE, intercurrent event; MAR, missing-at-random; MI, multiple imputation; SMN, survival motor neuron.

6.3. Multiplicity

The hierarchical testing approach will be used to control the overall significance level at $\alpha = 0.05$, which allows for multiple tests using $\alpha = 0.05$ at each step for the confirmatory tests. To control the multiplicity of the 2 hypotheses for the primary objectives in the same step, Step 1, the Hochberg procedure (Hochberg 1988) will be used to evaluate the 2 comparisons.

The order of the hierarchy is defined below (Step 1 to Step 8):

- **Step 1**. The 2 comparisons (Comparison 1 and Comparison 2) for the primary analysis of the primary efficacy endpoint, Change from Baseline in HFMSE total score at 12 months, will be evaluated simultaneously using the Hochberg procedure.
  - Comparison 1: The primary analysis of the main estimand for the primary efficacy endpoint, Change from Baseline in HFMSE total score at 12 months, compared between 20 mg/kg and placebo.
  - Comparison 2: The primary analysis of the main estimand for the primary efficacy endpoint, Change from Baseline in HFMSE total score at 12 months, compared between apitegromab combined dose and placebo.
- **Step 2**. The primary analysis of the key secondary efficacy endpoint, Change from Baseline in RULM total score at 12 months, compared between apitegromab 20 mg/kg and placebo.
- **Step 3**. The primary analysis of the key secondary efficacy endpoint, proportion of patients with ≥3-point change from Baseline in the HFMSE total score at 12 months, compared between apitegromab 20 mg/kg and placebo.
- **Step 4**. The primary analysis of the key secondary efficacy endpoint, Change from Baseline in number of WHO motor development milestones attained at 12 months, compared between apitegromab 20 mg/kg and placebo.
- **Step 5**. The primary analysis of the main estimand for the primary efficacy endpoint, Change from Baseline in HFMSE total score at 12 months, compared the difference between apitegromab 10 mg/kg and placebo.
- **Step 6**. The primary analysis of the key secondary efficacy endpoint, Change from Baseline in RULM total score at 12 months, compared between apitegromab 10 mg/kg and placebo.
- **Step 7**. The primary analysis of the key secondary efficacy endpoint, proportion of patients with ≥3-point change from Baseline in the HFMSE total score at 12 months, compared between apitegromab 10 mg/kg and placebo.
- **Step 8**. The primary analysis of the key secondary efficacy endpoint, Change from Baseline in number of WHO motor development milestones attained at 12 months, compared between apitegromab 10 mg/kg and placebo.

Only the primary analyses of the main estimands of the primary efficacy endpoint and the key secondary efficacy endpoints listed in the hierarchical procedure will be tested in a confirmatory way by order. These estimands are summarized in Table 8, Section 6.1, and Section 6.2.

If Step 1 is shown with a two-sided p-value ≤0.05 for both comparisons, both apitegromab 20 mg/kg and apitegromab combined dose achieved statistical significance for the primary endpoint, the hierarchical testing approach will continue to Step 2. A similar process would continue to the next step if the test in the previous step achieved statistical significance. If the test in a step is not shown with a p-value ≤0.05 to achieve the statistical significance, any tests in the remaining steps in the hierarchical procedure will be interpreted in an exploratory manner.

If Step 1 is shown one comparison with a p-value >0.05 and the other comparison with a p-value ≤0.025, the statistical significance is achieved for the primary endpoint with the comparison that has the p-value ≤0.025, but the hierarchical testing approach will stop, and tests specified in Step 2 to Step 8 will be interpreted in an exploratory manner.

If Step 1 is shown one comparison with a p-value >0.05 and the other comparison with a p-value >0.025, the statistical significance is not achieved for the primary endpoint. Tests specified in Step 2 to Step 8 will be interpreted in an exploratory manner.

6.4. Other and Additional Secondary Efficacy Endpoints

Any statistical comparison of other secondary efficacy endpoints will be performed in an exploratory manner.

6.4.1. HFMSE Total Score

The HFMSE total score and the change from Baseline will be summarized by treatment and by scheduled visit for the MITT Set, the PP Set, the EXP Set, and the Full Analysis Set. Similar summary tables will be repeated by type of SMN therapy at randomization (ie, nusinersen or risdiplam) and number of SMN therapies treated (1 or 2) prior to the study enrollment.

The number and percentage of patients achieving various magnitudes (such as ≥0, ≥1, ≥2, ≥3, ≥4, etc) of change from Baseline in HFMSE total score will be presented by treatment and scheduled visit on the MITT Set, the PP Set, the EXP Set, and the Full Analysis Set. Change from Baseline in HFMSE total score will be plotted by treatment and by scheduled visit on the MITT Set and the PP Set.

Statistical comparisons for the MEP or the Pooled Population will be performed in an exploratory manner, which are listed below.

6.4.1.1. Analyses for the Main Efficacy Population

In addition to the comparison at 12 months as the primary analysis of the primary endpoint described in Section 6.1.4, comparisons of change from Baseline in HFMSE total score for the MITT Set and the PP Set at other scheduled postbaseline visits will also be evaluated by the same analysis. The LS mean difference estimates at each visit (other than the 12-month visit) will be presented with the corresponding 95% CI and the p-value, respectively.

In addition to using the threshold of 3 points to categorize the change from Baseline in HFMSE total score (ie, <3 versus ≥3) at 12 months as the key secondary endpoint, the similar logistic regression model described in Section 6.2.2.4 will be performed using the observed cases for the MITT Set and the PP Set to evaluate the comparisons between each apitegromab treatment and placebo in the proportion of patients achieving various magnitudes (such as 0, 1, 2, 4, etc.) of

change from Baseline in HFMSE score at 12 months. The odds ratio, the respective 95% CI, and the p-value will be presented.

Time to achieve ≥3-point improvement from Baseline in HFMSE total score compared between each apitegromab dose (ie, 10 mg/kg or 20 mg/kg) and placebo will be summarized by Kaplan-Meier estimates for the MITT Set and the PP Set. The comparisons between each apitegromab dose and placebo will be evaluated by the log-rank test. The same analysis will also be conducted for comparison between apitegromab and placebo.

6.4.1.2. Analyses for the Exploratory Subpopulation

If there are sufficient (eg, at least 10) patients with evaluable values in change from Baseline in HFMSE total score at 12 months for each of the apitegromab 20 mg/kg and placebo groups, the statistical comparison between apitegromab 20 mg/kg and placebo for the EXP Set will be performed using the similar analysis for the primary endpoint described in Section 6.1.4 in which the model will include the fixed effects of treatment, visit, the treatment-by-visit interaction, Baseline HFMSE total score, Baseline HFMSE total score-by-time interaction, and type of SMN therapy at randomization (ie, nusinersen or risdiplam). The main and supportive estimands are defined similarly to those described in Section 6.1.1 and Section 6.1.2 but with the different population defined below. Any missing values will be handled in the same way as described in Section 6.1.3. Please note that this analysis will be under-powered.

Population: Patients in the EXP Set.

For binary variables using different thresholds to categorize the change from Baseline in HFMSE total score (such as <0 versus ≥0 and <3 versus ≥3, etc.) at 12 months, the similar logistic regression model described in Section 6.2.2.4 will be performed using observed cases for the EXP Set to evaluate the comparisons between apitegromab 20 mg/kg and placebo in the proportion of patients achieving various magnitudes (such as 0, 1, 2, 3, 4, etc.) of change from Baseline in HFMSE score at 12 months. The odds ratio, the respective 95% CI, and the p-value will be presented.

Time to decline (at least a -3-point change from Baseline in HFMSE total score) and time to therapeutic effect (ie, ≥1-point change from Baseline in HFMSE total score) compared between apitegromab 20 mg/kg and placebo will be summarized by Kaplan-Meier estimates for the EXP Set, respectively. If there are at least 10 evaluable values in change from Baseline in HFMSE total score at 12 months for each of the apitegromab 20 mg/kg and placebo groups, the comparisons between apitegromab 20 mg/kg and placebo will be evaluated by the log-rank test.

6.4.1.3. Analyses for the Main Efficacy Population/Exploratory Subpopulation Combined

The statistical comparison between apitegromab 20 mg/kg and placebo for patients from the Full Analysis Set will be performed using the same analysis for the primary endpoint described in Section 6.1.4. The main and supportive estimands are defined similarly to those described in Section 6.1.1 and Section 6.1.2 but with the different population defined below. Any missing values will be handled in the same way as described in Section 6.1.3 before analysis being conducted.

Population: Patients in the Full Analysis Set.

For binary variables using different threshold to categorize the change from Baseline in HFMSE total score (such as <0 versus ≥0 and <3 versus ≥3, etc.) at 12 months, the similar logistic regression model described in Section 6.2.2.4 will be performed using the observed cases for the Full Analysis Set to evaluate the comparisons between apitegromab 20 mg/kg and placebo in the proportion of patients achieving various magnitudes (such as 0, 1, 2, 3, 4, etc.) of change from Baseline in HFMSE score at 12 months. The odds ratio, the respective 95% CI, and the p-value will be presented.

Time to decline (at least a -3-point change from Baseline in HFMSE total score) and time to therapeutic effect (ie, ≥1-point change from Baseline in HFMSE total score) compared between apitegromab 20 mg/kg and placebo will be summarized by Kaplan-Meier estimates for the Full Analysis Set, respectively. The comparisons between apitegromab 20 mg/kg and placebo will be evaluated by the stratified log-rank test in which population (ie, MEP versus EXP) is considered as the stratification factor.

6.4.2. RULM Total Score

The RULM total score and the change from Baseline will be summarized by treatment and by scheduled visit for the MITT Set, the PP Set, the EXP Set, and the Full Analysis Set.

The number and percentage of patients achieving various magnitudes (such as ≥0, ≥1, ≥2, ≥3, ≥4, etc.) of change from Baseline in RULM total score will be presented by treatment and by scheduled visit for the MITT Set, the PP Set, the EXP Set, and the Full Analysis Set.

Statistical comparisons for the MEP, EXP or the Pooled Population will be performed in an exploratory manner, which are listed below.

6.4.2.1. Analyses for the Main Efficacy Population

In addition to the comparison at 12 months as primary analysis of the key secondary endpoint described in Section 6.2.1.4, comparisons of change from Baseline in RULM total score for the MITT Set and the PP Set at other scheduled postbaseline visits will also be evaluated by the same analysis. The LS mean difference estimates at each visit (other than the 12-month visit) will be presented with the corresponding 95% CI and the p-value, respectively.

The statistical comparison between each apitegromab dose (ie, 10 or 20 mg/kg) and placebo in the proportion of patients achieving various magnitudes (such as 0, 1, 2, 3, 4, etc.) of change at 12 months will be performed for the MITT Set and the PP Set using the logistic regression model with covariates of Baseline RULM total score, type of SMN therapy at randomization (ie, nusinersen or risdiplam), and age at initiation of SMN therapy (≥5 and <5 years). The odds ratio estimates will be presented with the corresponding 95% CIs and the p-values.

6.4.2.2. Analyses for the Exploratory Subpopulation

If there are at least 10 evaluable values in change from Baseline in RULM total score at 12 months for each of the apitegromab 20 mg/kg and placebo groups, the statistical comparison between apitegromab 20 mg/kg and placebo for the EXP Set will be performed using the same analysis for the key secondary endpoint described in Section 6.2.1.4. The main and supportive estimands are defined similarly to those described in Section 6.2.1.1 and Section 6.2.1.2 but with the different population defined below. Any missing values will be handled in the same way as described in Section 6.2.1.3.

Population: Patients in the EXP Set.

The statistical comparison between apitegromab 20 mg/kg and placebo in the proportion of patients achieving various magnitudes (such as 0, 1, 2, 3,4, etc.) of change at 12 months will be performed for the EXP Set using the logistic regression model with covariates of Baseline RULM total score and type of SMN therapy at randomization (ie, nusinersen or risdiplam). The odds ratio estimates will be presented with the corresponding 95% CIs and the p-values.

6.4.2.3. Analyses for the Main Efficacy Population/Exploratory Subpopulation Combined

The statistical comparison between apitegromab 20 mg/kg and placebo for patients from the Full Analysis Set will be performed using the same analysis for the primary endpoint described in Section 6.2.1.4. The main and supportive estimands are defined similarly to those described in Section 6.2.1.1 and Section 6.2.1.2 but with the different population defined below. Any missing values will be handled in the same way as described in Section 6.2.1.3.

Population: Patients in the Full Analysis Set.

The statistical comparison between apitegromab 20 mg/kg and placebo in the proportion of patients achieving various magnitudes (such as 0, 1, 2, 3, 4, etc.) of change at 12 months will be performed for the Full Analysis Set using the logistic regression model with covariates of Baseline RULM total score, type of SMN therapy at randomization (ie, nusinersen or risdiplam), and age at initiation of SMN therapy (≥5 and <5 years). The odds ratio estimates will be presented with the corresponding 95% CIs and the p-values.

6.4.3. WHO Motor Development Milestones

The number of WHO motor development milestones attained and the change from Baseline will be summarized by treatment and by scheduled visit for the MITT Set, the PP Set, the EXP Set, and the Full Analysis Set.

The number and percentage of patients attaining a new WHO motor development milestone relative to Baseline will be presented by treatment and by scheduled visit for the MITT Set, the PP Set, the EXP Set, and the Full Analysis Set.

Statistical comparisons for the MEP or the Pooled Population will be performed in an exploratory manner, which are listed below.

6.4.3.1. Analyses for the Main Efficacy Population

In addition to the comparison at 12 months as primary analysis of the key secondary endpoint described in Section 6.2.3.4, comparisons of change from Baseline in number of WHO motor development milestones attained for the MITT Set and the PP Set at other scheduled postbaseline visits will also be evaluated by the same analysis. The LS mean difference estimates at each visit (other than the 12-month visit) will be presented with the corresponding 95% CI and the p-value, respectively.

The statistical comparison between each apitegromab treatment (ie, 10 mg/kg or 20 mg/kg) and placebo in the proportion of patients who attain a new WHO motor development milestone relative to Baseline at 12 months will be performed for the MITT Set and the PP Set using the logistic regression model with covariates of Baseline RULM total score, type of SMN therapy at

randomization (ie, nusinersen or risdiplam), and age at initiation of SMN therapy (≥5 and <5 years). The odds ratio estimates will be presented with the corresponding 95% CIs and the p-values.

6.4.3.2. Analyses for the Exploratory Subpopulation

If there are at least 10 evaluable values in change from Baseline in number of WHO motor development milestones attained at 12 months for each of the apitegromab 20 mg/kg and placebo groups, the statistical comparison between apitegromab 20 mg/kg and placebo for the EXP Set will be performed using the same analysis for the key secondary endpoint described in Section 6.2.3.4. The main and supportive estimands are defined similarly to those described in Section 6.2.3.1 and Section 6.2.3.2 but with the different population defined below. Any missing values will be handled in the same way as described in Section 6.2.3.3.

Population: Patients in the EXP Set.

The statistical comparison between apitegromab 20 mg/kg and placebo in the proportion of patients who attain a new WHO motor development milestone relative to Baseline at 12 months will be performed for the EXP Set using the logistic regression model with covariates of the number of WHO motor development milestones attained at Baseline and type of SMN therapy at randomization (ie, nusinersen or risdiplam). The odds ratio estimates will be presented with the corresponding 95% CIs and the p-values.

6.4.3.3. Analyses for the Main Efficacy Population/Exploratory Subpopulation Combined

The statistical comparison between apitegromab 20 mg/kg and placebo for patients from Full Analysis Set will be performed using the same analysis for the primary endpoint described in Section 6.2.3.4. The main and supportive estimands are defined similarly to those described in Section 6.2.3.1 and Section 6.2.3.2 but with the different population defined below. Any missing values will be handled in the same way as described in Section 6.2.3.3.

Population: Patients in the Full Analysis Set.

The statistical comparison between apitegromab 20 mg/kg and placebo in the proportion of patients who attain a new WHO motor development milestone relative to Baseline at 12 months will be performed for the Full Analysis Set using the logistic regression model with covariates of the number of WHO motor development milestones attained at Baseline, type of SMN therapy at randomization (ie, nusinersen or risdiplam), and age at initiation of SMN therapy (≥5 and <5 years). The odds ratio estimates will be presented with the corresponding 95% CIs and the p-values.

6.4.4. Pediatric Evaluation of Disability Inventory Computer Adaptive Test

The PEDI-CAT score for each domain (ie, Mobility domain and Daily Activities domain) and the corresponding change from Baseline will be summarized by treatment and by scheduled visit for the MITT Set, EXP Set, and Full Analysis Set.

The relevant listing of PEDI-CAT at the patient level will be provided.

6.4.5. PROMIS Fatigue Questionnaire

The PROMIS score and the change from Baseline will be summarized by type of assessment (ie, adult form assessments, pediatric form assessments, and parent proxy assessments), by treatment, and by scheduled visit for the MITT Set, EXP Set, and Full Analysis Set.

The relevant listing of PROMIS scores at the patient level will be provided.

6.4.6. Assessment of Caregiver Experience with Neuromuscular Disease

The ACEND total score, 2 domain scores and 7 subdomain scores, and the change from Baseline will be summarized by treatment and by scheduled visit for the MITT Set, EXP Set, and Full Analysis Set.

The relevant listing of ACEND at the patient level will be provided.

6.5. Subgroup Analysis

The consistency of the treatment effect for the efficacy endpoints of HFMSE total score and RULM total score will be explored for the following: type of SMN therapy at randomization, age at initiation of SMN therapy, and region:

- Type of SMN therapy at randomization (ie, nusinersen or risdiplam)
- Age at initiation of SMN therapy (≥5 and <5 years)
- Region (Europe or North America)

For the type of SMN therapy at randomization and the age at initiation of SMN therapy, an additional interaction term of treatment by subgroup will be added into the model of the primary analysis. For the region subgroup, the additional subgroup and treatment-by-subgroup terms will be added into the model of the primary analysis. P-values will be interpreted in an exploratory manner.

Within each subgroup, if there are sufficient (eg, at least 10) patients with evaluable values, primary analysis specified in Section 6.1.4 and Section 6.2.2.4 will be conducted.

7. SAFETY ANALYSIS

The Safety Set will be used for safety analyses. Safety will be assessed based on AEs, clinical laboratory data, vital signs, electrocardiogram (ECG) parameters, C-SSRS, and physical examinations.

7.1. Adverse Events/Adverse Drug Reactions

All AEs, including SAEs reported will be assessed by the Investigator according to the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) grading scale. All AEs will be coded by SOC and PT using Medical Dictionary for Regulatory Activities (MedDRA).

All AEs will be analyzed based on the principle of treatment emergence. Treatment-emergent adverse events (TEAEs), including SAEs (ie, serious TEAEs), will be summarized in the safety analyses unless otherwise specified. A TEAE is defined as an AE that started or worsened in severity after the start of the first dose of study drug. Adverse events with missing or partial start dates will be considered as treatment emergent unless the partial date or the reported end date excludes that possibility, such as the AE month is prior to the study drug administration month. Adverse events starting on the date of the first dose will be considered treatment emergent unless the AE onset time is recorded and is prior to the start of the infusion.

For summaries of TEAEs by SOC and PT, a patient will be counted once at the SOC level and once at each PT within the SOC level. For summaries of TEAE by SOC, PT, and maximum CTCAE grade, a patient will be counted once at the highest CTCAE grade level for which the event occurred at the SOC level and the highest CTCAE grade level for each unique PT within that SOC level. In cases when CTCAE is missing for a TEAE, a Grade 3 will be imputed. In cases when the relationship to study drugs is missing for a TEAE, it will be considered as related to study drugs.

The summaries presenting frequency of TEAEs by SOC and PT will be ordered by overall descending frequency of SOC and then, within an SOC, by overall descending frequency of PT.

An overall summary table will be provided, which includes the number of TEAEs, the number and percentage of patients will be presented for TEAEs, serious TEAEs, study drug-related TEAEs, study drug-related serious TEAEs, SMN therapy-related TEAE, SMN therapy-related serious TEAE, procedure-related TEAEs, Grade 3+ TEAEs, TEAEs associated with the underlying disease and/or disease progression, time to onset (≤12 weeks or >12 weeks), TEAEs leading to discontinuation of study drug, TEAEs leading to drug interruption, TEAEs leading to dose modification, and TEAEs leading to death. The time to onset is defined as the length of time from the first dose of study drug to the AE onset.

The overall summary table will be repeated by type of SMN therapy at randomization (ie, nusinersen or risdiplam).

The following AE tables will be presented by treatment and by patient population (ie, MEP, EXP, and Pooled Population):

- Summary of TEAEs by SOC and PT
- Summary of TEAEs by SOC and PT with at least 5% PT

- Summary of TEAE that occurred in at least 5% of apitegromab patients and occurred at least 5% higher in incidence for apitegromab compared to placebo by SOC and PT
- Summary of TEAEs by SOC and PT, and Maximum severity
- Summary of TEAEs by PT
- Summary of Study Drug-related TEAEs by SOC and PT
- Summary of Study Drug-related TEAEs by SOC and PT, and Maximum severity
- Summary of Serious TEAEs by SOC and PT
- Summary of Serious TEAEs by SOC and PT, and Maximum severity
- Summary of Study Drug-related Serious TEAEs by SOC and PT
- Summary of Study Drug-related Serious TEAEs by SOC and PT, and Maximum severity
- Summary of Procedure-related TEAEs by SOC and PT
- Summary of Grade 3+ TEAEs by SOC and PT
- Summary of Study Drug-related Grade 3+ TEAEs by SOC and PT
- TEAEs leading to discontinuation of study drug by SOC and PT
- TEAEs leading to dose modification or interruptions by SOC and PT
- TEAEs leading to withdrawal from study by SOC and PT
- TEAEs leading to death by SOC and PT
- Summary of TEAE by Standardized MedDRA Queries (SMQ) and PT

Similar AE tables will be repeated by type of SMN therapy at randomization (ie, nusinersen or risdiplam) and by time to onset (≤12 weeks or >12 weeks), respectively. When 20 or more TEAEs associated with the underlying disease and/or disease progression are reported in the study, similar AE tables will also be repeated by association (yes or no) with the underlying disease and/or disease progression.

The following AE listings will be provided, which include those that are not TEAEs:

- All AEs
- Grade 3+ AEs
- Serious AEs
- Treatment-related SAEs
- AEs associated with the underlying disease and/or disease progression
- AEs leading to permanent discontinuation of study drug
- AEs leading to withdrawal from study
- All deaths

Any pregnancy will be captured as an AE, and a listing of pregnancy-associated AEs will be provided if any pregnancy occurred.

7.2. Laboratory Evaluations

The parameters from hematology, chemistry, urinalysis and coagulation listed below will be assessed. The laboratory abnormality categories (low, normal, or high) will be determined based on the corresponding age-associated normal reference range for each parameter provided in the safety laboratory manual.

- Hematology: absolute platelet count, erythrocytes (ie, red blood cell [RBC]) count, hemoglobin, hematocrit, mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH), MCH concentration, %reticulocytes, leukocytes (ie, white blood cell [WBC]) count, neutrophils count, neutrophils/leukocytes (%), lymphocytes count, lymphocytes/leukocytes (%), monocytes count, monocytes/leukocytes (%), eosinophils, and basophils
- Chemistry: blood urea nitrogen (BUN), potassium, aspartate aminotransferase (AST), total and direct bilirubin (BILI), creatinine, sodium, alanine aminotransferase (ALT), total protein, glucose (nonfasting), calcium, alkaline phosphatase (ALP), albumin, carbon dioxide, chloride, creatine phosphokinase, gamma-glutamyl transferase (GGT), lactate dehydrogenase (LDH), magnesium, phosphate, total cholesterol, triglycerides, uric acid
- Urinalysis: specific gravity, pH, glucose, protein, blood, ketones, BILI, nitrite
- Coagulation: activated partial thromboplastin time (aPTT), prothrombin time/international normalized ratio (PT/INR).

All quantitative clinical laboratory results of hematology, chemistry, urinalysis, and coagulation parameters will be summarized. The quantitative clinical laboratory results and the corresponding change from baseline will be plotted. For urinalysis parameters, any qualitative assessments will be summarized by visit.

Laboratory abnormalities will be summarized with shift from baseline tables. Each hematology and chemistry value will be flagged as “low”, “normal”, or “high” relative to the age based normal reference ranges, or as “unknown” if no results are available. Each urinalysis value will be flagged as “positive”, “negative”, or “unknown” if no values are available. Shifts from baseline to high/low status for hematology and chemistry parameters and shifts from baseline to high/positive for urinalysis will be presented. In each summary, the denominator for the percentage is the number of patients at risk for the shift. The number at risk for a shift to low is the number of subjects whose baseline value was not low and who had at least one postbaseline value. The number at risk for a shift to high is the number of subjects whose baseline value was not high and who had at least one postbaseline value. Subjects will be counted only once for each parameter and each type of shift regardless of how many postbaseline assessments had that type of shift. A shift to high includes normal to high, low to high, and unknown to high; a shift to low includes normal to low, high to low, and unknown to low. A shift to positive includes ‘negative’ to ‘positive’ and ‘unknown’ to ‘positive’.

The shift from baseline in CTCAE will also be summarized for an increase to any CTCAE grade, that is, increase to Grade 3 or Grade 4.

Abnormal platelet count will be evaluated by different cutoff levels, and the number and percentage of subjects will be summarized.

The drug-induced liver injury (DILI) evaluation will be summarized by the number and percentage of patients who meet any of the following criteria: ALT >3× upper limit of normal (ULN), ALT>5× ULN, ALT>10× ULN, AST >3× ULN, AST >5× ULN, AST >10× ULN, ALP >1.5× ULN, BILI >1.5× ULN, BILI >2× ULN, ALT or AST >3× ULN, ALT or AST >5× ULN, ALT or AST >10× ULN, ALT or AST >3× ULN and BILI >1.5× ULN, and ALT or AST >3× ULN and BILI >2× ULN.

If any patient meets the Hy's Law, a plot will be provided.

Individual results that are outside of normal reference ranges will be flagged in data listings.

7.3. Vital Signs

Vital signs and physical characteristics (ie, heart rate [beats/min], systolic blood pressure [mmHg], diastolic blood pressure [mmHg], respiratory rate [breaths/min], height [cm], weight [kg], weight percentile [%], BMI [kg/m^2], and BMI percentile [%]) will be summarized and plotted. The vital signs are measured at multiple time points at each schedule visit, and the average of nonmissing results at each postbaseline visit will be used in the summary where applicable.

The incidence of abnormal vital signs will be provided for abnormal vital signs determined by each criterion listed in Table 11.

Table 11: Criteria to Determine the Abnormal Vital Signs

Vital Signs	Criteria
Systolic Blood Pressure (mmHg)	• <90 mmHg • >140 mmHg • >160 mmHg • Increment >20 mmHg • Increment >40 mmHg • Decrement >20 mmHg • Decrement >40 mmHg
Diastolic Blood Pressure (mmHg)	• <50 mmHg • >90 mmHg • >100 mmHg • Increment >10 mmHg • Increment >20 mmHg • Decrement >10 mmHg • Decrement >20 mmHg
Weight (kg)	• Decrease of ≥7% from baseline • Increase of ≥7% from baseline
Temperature (°C)	• >38.0°C • <36.0°C
Respiratory Rate (breaths/min)	• <12 breaths/min • >20 breaths/min • Increment >10 breaths/min • Increment >20 breaths/min • Decrement >10 breaths/min • Decrement >20 breaths/min
Heart Rate (beats/min)	• <60 beats/min • >100 beats/min • Increment >20 beats/min • Increment >40 beats/min • Decrement >20 beats/min • Decrement >40 beats/min

7.4. Electrocardiogram

Electrocardiograms (ECGs) are performed in triplicates, and the average of the ECG triplicates will be used for analysis of ECG quantitative assessments, unless otherwise specified.

The ECG quantitative assessments and the corresponding change from baseline for each ECG parameter (heart rate [bpm], PR Interval [msec], RR Interval [msec], QRS duration (msec), QT interval [msec], Corrected QT Bazett [QTcB] interval [msec], and Corrected QT Fridericia [QTcF] interval [msec]) will be summarized.

Abnormal ECG quantitative assessments will be determined by the age-associated normal reference ranges listed in Table 12. ECG abnormality for quantitative assessments will be summarized by shift tables.

Table 12: Criteria to Determine Abnormal ECG Quantitative Assessments

ECG Assessments	Criteria
PR Interval (msec)	• >150 (msec) for patients with age <3 years • >180 (msec) for patients with age 3 to 11 years • >200 (msec) for patients with age ≥12 years
QTcB interval (msec)	• >449 (msec) for patients with age <18 years • >499 (msec) for patients with age ≥18 years
QTcF interval (msec)	• >449 (msec) for patients with age <18 years • >499 (msec) for patients with age ≥18 years
QRS duration (msec)	• >79 (msec) for patients with age <3 years • >89 (msec) for patients with age 3 to 11 years • >99 (msec) for patients with age 12 to 17 years • >109 (msec) for patients with age ≥18 years

Abbreviations: ECG, electrocardiogram; QTcB interval, QT interval corrected by Bazett's formula; QTcF interval, QT interval corrected by Fridericia's formula.

In addition, the number and percentage of patients with corrected QT (QTc) interval that meet the criteria listed below will be summarized.

- >450 msec
- >480 msec
- >500 msec
- Increase from baseline >30 msec
- Increase from baseline >60 msec
- >500 msec and >60 msec increase from baseline

The ECG results interpretation will be summarized.

7.5. Columbia-Suicide Severity Rating Scale

The number of patients who performed the C-SSRS and the number and percentage of patients with positive suicidal ideation and suicidal behavior will be presented by treatment and by visit. The treatment-emergent suicidal ideation and suicidal behavior will be summarized by treatment. In addition, the severity of suicidal ideation and the corresponding change from Baseline will be

summarized by treatment and by visit. The Safety Set will be evaluated for C-SSRS; a similar analysis will also be evaluated by each SMN therapy.

The relevant listing of C-SSRS scores and each domain at the patient level will be provided.

7.6. Physical Examination

Physical examination data will only be provided in a data listing.

8. PHARMACOKINETICS

PK measurement of serum concentrations will be summarized for the PK Analysis Set using descriptive statistics at the scheduled visits. The number of evaluable patients, arithmetic mean, SD, median, 25% and 75% quartiles, minimum, maximum, geometric mean, and coefficient of variation (CV) will be presented at each visit.

The mean serum concentrations (±SD) will be plotted over time by treatment on the linear and logarithmic-linear scales. Atypical drug concentrations (eg, very low or very high) will be excluded from the analysis, if no apparent explanation exists. Concentration observations will also be removed from the dataset if corresponding dosing or sampling times are missing or cannot be reconstructed. Measurements that are below the level of quantification (BLQ) will be set to missing and removed from the analysis before statistics are calculated, and if the BLQ occurs before the first measurable concentration (such as predose), it should be set to zero. If the value of the concentration is "Non-Determinable", then the concentration value will be set to missing.

A listing of individual serum concentration data will be provided.

Individual PK concentration will be plotted for patients who received either apitegromab 10 mg/kg or 20 mg/kg on a linear and a semi-log scale against the actual study days of collecting the PK samples.

9. PHARMACODYNAMICS

The circulating serum total latent myostatin concentration and the change from Baseline will be summarized by treatment and by scheduled visit for the PD Analysis Set. A relevant listing of the total latent myostatin concentrations at the patient level will be provided.

The apitegromab and total latent myostatin concentrations from this trial may be combined with other trial data to support the population PK analysis, PK/PD, and exposure-response analysis. The planned analyses will be outlined in a PK-PD and exposure-response analysis plan and documented in a separate report.

10. IMMUNOGENICITY

Serum samples will be collected to evaluate the presence or absence of antibodies to apitegromab in serum.

Serum samples will be screened for antibodies binding to apitegromab using a validated electrochemiluminescence (ECL) method. For samples that test positive in the confirmatory assay, titers will be determined, and the samples will be tested using a validated neutralization antibody assay.

Antidrug antibody results will be presented by treatment using the Immunogenicity Analysis Set. The incidence of antibodies to apitegromab will be summarized over time by treatment using descriptive statistics. The incidence of treatment-emergent positive ADA result will also be summarized where the treatment-induced positive and treatment-boosted positive are defined as follows, respectively.

- Treatment-induced positive: baseline ADA is missing or negative, and subject has any post-treatment positive ADA sample based on results of confirmatory testing.
- Treatment-boosted positive: baseline ADA is positive based on results from confirmatory testing and higher levels of ADA titer in any post-treatment sample (ie, any post-baseline ADA titer is ≥4-times the baseline ADA titer).

A relevant listing of immunogenicity results at the patient level will also be provided.

11. INTERIM ANALYSIS

No interim analysis of efficacy will be conducted for this study.

12. CHANGES FROM ANALYSIS PLANNED IN PROTOCOL

Not applicable

13. REFERENCES

Ader D. Developing the Patient-Reported Outcomes Measurement Information System (PROMIS). *Med Care*. 2007;5(5):S1-S2.

Conover, W. J., and Ronald L. Iman. "Analysis of Covariance Using the Rank Transformation." *Biometrics*, vol. 38, no. 3, 1982, pp. 715–24. *JSTOR*, https://doi.org/10.2307/2530051.

Conway PM, Erlangsen A, Teasdale TW, et al. Predictive validity of the Columbia-Suicide Severity Rating Scale for short-term suicidal behavior: a Danish study of adolescents at a high risk of suicide. *Arch Suicide Res*. 2016;21(3):455-469.

Gauld LM, Kappers J, Carlin JB, Robertson CF. Height prediction from ulna length. *Dev Med Child Neurol*. 2004(Jul);46(7):475-80. doi: 10.1017/s0012162204000787. PMID: 15230461.

Gipson PY, Agarwala P, Opperman KJ, et al. Columbia-Suicide Severity Rating Scale. *Pediatr Emerg Care*. 2015;31(2):88-94.

Glanzman AM, O'Hagen JM, McDermott MP, et al. Validation of the Expanded Hammersmith Functional Motor Scale in spinal muscular atrophy type II and III. J Child Neurol. 2011;26(12):1499-1507.

Haley SM, Raczek AE, Coster WJ, et al. Assessing mobility in children using a computer adaptive testing version of the pediatric evaluation of disability inventory. *Arch Phys Med Rehabil*. 2005;86(5):932-939.

Hochberg, Y. A Sharper Bonferroni Procedure for Multiple Significance Testing. *Biometrika*. 1988;75:800–803.

ICH E9 (R1): Addendum on Estimands and Sensitivity Analysis in Clinical Trials to the Guideline on Statistical Principles for Clinical Trials. 2019.

Kowalewski EK, Weideman AMK, Koch GG. SAS Macro for Randomization-Based Methods for Covariance and Stratified Adjustment of Win Ratios and Win Odds for Ordinal Outcomes. *SESUG Paper. 139-2023.*

Li KH. Imputation Using Markov Chains. *J Stat Comput Simu*. 1988;30:57-79.

Little, RJA. Pattern-mixture models for multivariate incomplete data. *J Am Stat Assoc*. 1993;88:125-134.

Mantel N. Chi-Square Tests with One Degree of Freedom, Extensions of the Mantel–Haenszel Procedure. *J Am Stat Assoc*. 1963(September);58 (303):690-700.

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):284-292.

Mazzone ES, Mayhew A, Montes J, et al. Revised upper limb module for spinal muscular atrophy: Development of a new module. *Muscle Nerve*. 2017;55(6):869-874.

Nilsson ME, et al. Columbia-suicide severity rating scale scoring and data analysis guide. C-SSRS Scoring Version 2. 2013:1-13.

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):693-697.

Pasternak A, Sideridis G, Fragala-Pinkham M, et al. Rasch analysis of the Pediatric Evaluation of Disability Inventory-computer adaptive test (PEDI-CAT) item bank for children and young adults with spinal muscular atrophy. *Muscle Nerve*. 2016;54(6):1097-1107.

Posner K, Brown GK, Stanley B, et al. The Columbia-Suicide Severity Rating Scale: initial validity and internal consistency findings from 3 multisite studies with adolescents and adults. *Am J Psychiatry*. 2011;168(12):1266-1277.

Ratitch B, O'Kelly M. Implementation of Pattern-Mixture Models Using Standard SAS/STAT Procedures. In Proceedings of PharmaSUG 2011 (Pharmaceutical Industry SAS Users Group). Paper SP04. Cary, NC: SAS Institute Inc. 2011.

Ratitch B, Lipkovich I, O'Kelly M. Combining Analysis Results From Multiply Imputed Categorical Data. In Proceedings of PharmaSUG 2013 (Pharmaceutical Industry SAS Users Group). Paper SP03. Cary, NC: SAS Institute Inc. 2013.

Rubin DB. Multiple Imputation for Nonresponse in Surveys. John Wiley & Sons. 1987.

Schafer JL. Analysis of Incomplete Multivariate Data. Chapman and Hall, New York. 1997.

WHO Multicentre Growth Reference Study Group. WHO Motor Development Study: windows of achievement for 6 gross motor development milestones. *Acta Paediatr Suppl*. 2006;450:86-95.

Weideman AMK, Kowalewski EK, Koch GG. Randomization Based Methods for Covariance and Stratified Adjustment of Win Ratios and Win Odds for Ordinal Outcomes. *JSM 2022 – Biopharmaceutical Section*.

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-S45.

Young SD, Montes J, Salazar R, et al. Scoliosis surgery significantly impacts motor abilities in higher-functioning individuals with spinal muscular trophy. *J Neuromuscul Dis*. 2020;7(2):183-192.

14. APPENDICES

14.1. Appendix A: Schedule of Activities/Assessments for Study SRK-015-003

Activity/Assessment	**SCR**	**Treatment Period**														**Follow-up**		
Visit **Time Point (Study Day)**	**SCR** **-28 to -1**	**V1** **1**	**V2** **29**	**V3** **57**	**V4** **85**	**V5** **113**	**V6** **141**	**V7** **169**	**V8** **197**	**V9** **225**	**V10** **253**	**V11** **281**	**V12** **309**	**V13** **337**	**V14** **365/ EOS/EOT**	**Unsch**	**V15** **393**	**V16** **421**	**V17** **505/ EOS/ET**
Visit window (±days)			**±7**	**±7**	**±7**	**±7**	**±7**	**±7**	**±7**	**±7**	**±7**	**±7**	**±7**	**±7**	**±7**		**±7**	**±7**	**±7**
Informed Consent	X																		
Demographics and Disease/Medical History	X																		
Inclusion/Exclusion	X																		
Pregnancy Test (if applicable)	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Weight	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Height	X	X		X		X		X		X		X			X	X	X	X	X
Physical Examination	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Vital Signs	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Clinical Safety Laboratory Assessments	X	X	X	X		X		X		X		X		X	X	X	X	X	X
12-lead ECG	X	X	X			X		X		X		X		X	X	X			X
PK and PD Sampling		X	X			X		X		X		X		X	X	X			X
ADA Sampling		X	X			X		X		X		X		X	X	X			X
Randomization		X																	
Study Drug Administration		X	X	X	X	X	X	X	X	X	X	X	X	X					
Motor Function Outcome Measures	X	X		X		X		X		X		X			X	X	X	X	X

Activity/Assessment	**SCR**	**Treatment Period**															**Follow-up**		
Visit **Time Point (Study Day)**	**SCR** **-28 to -1**	**V1** **1**	**V2** **29**	**V3** **57**	**V4** **85**	**V5** **113**	**V6** **141**	**V7** **169**	**V8** **197**	**V9** **225**	**V10** **253**	**V11** **281**	**V12** **309**	**V13** **337**	**V14** **365/ EOS/EOT**	**Unsch**	**V15** **393**	**V16** **421**	**V17** **505/ EOS/ET**
Visit window (±days)			**±7**	**±7**	**±7**	**±7**	**±7**	**±7**	**±7**	**±7**	**±7**	**±7**	**±7**	**±7**	**±7**		**±7**	**±7**	**±7**
PEDI-CAT and PROMIS		X		X		X		X		X		X			X	X	X	X	X
ACEND Questionnaire		X						X							X	X		X	X
C-SSRS		X						X							X	X		X	X
Site Check-in		X	X	X	X	X	X	X	X	X	X	X	X	X					
Adverse Event Recording	To be collected from the date the Informed Consent Form (ICF) is signed through the last trial visit																		
SAE Reporting	To be reported from the date the ICF is signed through the last trial visit																		
Concomitant Treatment Recording	To be collected from the date the ICF is signed through the last trial visit																		

ACEND: Assessment of Caregiver Experience with Neuromuscular Disease; ADA: antidrug antibody; C-SSRS: Columbia-Suicide Severity Rating Scale; ECG: electrocardiogram; EOS: end of study; EOT: end of treatment; ET: early termination; ICF: informed consent form; PD: pharmacodynamic; PEDI-CAT: Pediatric Evaluation of Disability Inventory Computer Adaptive Test; PK: pharmacokinetic; PROMIS: Patient-Reported Outcomes Measurement Information System;: serious adverse event; SCR: Screening; Unsch: unscheduled; V: visit.

14.2. Appendix B: Sample SAS Code Applying the MI Under the MNAR Assumption

```
/* ---     Step 1: Impute the intermediate missing to make the data monotone missing     --- */
PROC MI DATA=df1 nimpute=100 out=df2 seed=1053;
BY trt;
VAR V1 V3 V5 V7 V9 V11 V14;
MCMC CHAIN=multiple IMPUTE=monotone;
RUN;
/*---     Step 2: impute the data with monotone missing using MI under MNAR, assuming trt=1 is placebo ---*/
PROC SORT DATA=df2; BY _IMPUTATION_; RUN;
PROC MI DATA =df2 NIMPUTE=1 OUT=df3 seed=1053;
BY _IMPUTATION_;
CLASS TRT;
VAR V1 V3 V5 V7 V9 V11 V14;
MONOTONE reg(/details);
MNAR model (V3 V5 V7 V9 V11 V14/ MODELOBS = (trt = '1'));
RUN;
```

14.3. Appendix C: Sample SAS Code Applying the MI Under the MAR Assumption

```
/* ---  Step 1: Impute the intermediate missing to make the data monotone missing --- */
PROC MI DATA=df1 nimpute=100 out=df2 seed=1053;
BY trt;
VAR V1 V3 V5 V7 V9 V11 V14;
MCMC CHAIN=multiple IMPUTE=monotone;
RUN;
/*---  Step 2: impute the data with monotone missing using MI under MAR. ---*/
PROC SORT DATA=df2; BY _IMPUTATION_; RUN;
PROC MI DATA =df2 NIMPUTE=1 OUT=df3 seed=1053;
BY _IMPUTATION_;
CLASS trt;
VAR trt V1 V3 V5 V7 V9 V11 V14;
MONOTONE reg(/details);
RUN;
```

14.4. Appendix D: Sample SAS Code Performing the MMRM on Imputed Datasets

```
/*** assume 7 timepoints ***/

PROC MIXED data=df3;

by _IMPUTATION_;

CLASS trt visit subjid strvar1 strvar2;

MODEL change= trt visit trt*visit Baseline Baseline*visit strvar1 strvar2/
DDFM=KenwardRoger;

REPEATED visit / TYPE=un SUBJECT=subjid;

LSMESTIMATE trt*visit 'pbo at timepoint 7' 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0;

LSMESTIMATE trt*visit 'low at timepoint 7' 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0;

LSMESTIMATE trt*visit 'high at timepoint 7' 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1;

LSMESTIMATE trt*visit 'active at timepoint 7' 0 0 0 0 0 0 0 0 0 0 0 0 0.5 0 0 0 0 0 0 0.5;

ESTIMATE "low vs. pbo at timepoint 7"
        trt -1 1 0    trt*visit 0 0 0 0 0 0 -1 0 0 0 0 0 0 1 0 0 0 0 0 0;

ESTIMATE "high vs. pbo at timepoint 7"
        trt -1 0 1    trt*visit 0 0 0 0 0 0 -1 0 0 0 0 0 0 0 0 0 0 0 0 1;

ESTIMATE "active vs. pbo at timepoint 7"
        trt -1 0.5 0.5    trt*visit 0 0 0 0 0 0 -1 0 0 0 0 0 0 0.5 0 0 0 0 0 0 0.5;

ODS OUTPUT ESTIMATES=LSEST LSMEstimates=LSM;

RUN;

/* use Rubin's rule*/

PROC SORT DATA = LSEST; BY LABEL; run;

PROC MIANALYZE DATA=LSEST;

BY LABEL;

MODELEFFECTS Estimate;

STDERR stderr;

ODS OUTPUT PARAMETERESTIMATES=LSEST_RESULT;

RUN;


PROC SORT DATA =LSM; BY LABEL; run;
```

```
PROC MIANALYZE DATA =LSM;

BY LABEL;

MODELEFFECTS Estimate;

STDERR stderr;

ODS OUTPUT PARAMETERESTIMATES=LSM_RESULT;

RUN;
```

14.5. Appendix E: Sample SAS Code Performing the Randomization-Based Method for Covariance and Stratified Adjustment of Win Odds for Ordinal Outcome on Imputed Datasets

SAS Macro of the randomization based method for covariance and stratified adjustment of winodds for ordinal outcome is available on GitHub

https://github.com/elaineek/adj-wrwo/blob/main/Adj_WinOdds.sas

```
/*** Step 1: generate rank data based on continuous variable within each stratum ***/
proc sort data=dat; by strvar1_2;  run;
proc rank data=dat out=dat_rank ties=dense;
by strvar1_2;
var v1 chg2-chg5;
ranks baseRank chgrankV2 chgrankV3 chgrankV4 chgrankV5;
run;
/*** Step 2: perform the randomization based method for covariance and stratified adjustment of winodds for ordinal outcome is available ***/
%Adj_WinOdds(DSNIN = dat_rank,
DSNOUT =out,
PID = subjid,
OUTCOMES =chgrankV2 chgrankV3 chgrankV4 chgrankV5,
ARM = trt,
baseline = v1,
strata= strvar1_2);
/*** Step 3: combine results and apply Rubin's rules ***/
PROC MIANALYZE data=result;
BY visit;
MODELEFFECTS logWO;
stderr SE_logWO;
ODS OUTPUT PARAMETERESTIMATES=PARAMS;
RUN;
```

14.6. Appendix F: Sample SAS Code Performing the ANCOVA on Imputed Datasets

```
PROC GLM DATA=df;
by IMPUTEN;
CLASS trt strvar1 strvar2;
MODEL change = trt strvar1 strvar2 Baseline/ solution;
LSMEANS trt/ stderr pdiff cov;
estimate 'active vs pbo at last timepoint' trt -1 0.5 0.5;
estimate 'high vs pbo at last timepoint' trt -1 0 1;
estimate 'low vs pbo at last timepoint' trt -1 1 0;
/* assuming the mean Baseline is 50 from the dataset of df */
estimate 'active at last timepoint' intercept 1 trt 0 0.5 0.5 Baseline 50;
ods output LSMeans=LSMeans; estimates=est_ancova;
RUN;
/* LS Mean (SE) from ANCOVA for each group */
proc sort data=LSMeans; by effect trt; run;
PROC MIANALYZE data=LSMeans;
BY effect trt;
MODELEFFECTS LSMean;
stderr stderr;
ODS OUTPUT PARAMETERESTIMATES=result;
RUN;
/* LS Mean difference from ANCOVA for comparing to pbo */
proc sort data=est_ancova; BY parameter; run;
PROC MIANALYZE data=est_ancova;
BY parameter;
MODELEFFECTS Estimate;
stderr stderr;
ODS OUTPUT PARAMETERESTIMATES=result_diff_ancova;
RUN;
```

14.7. Appendix G: Details of Performing the Tipping-point Analysis

The tipping point analysis assumes MAR in the placebo group and MNAR in the active treatment groups and it imputes the missing values for patients in active drug group by adding a shift value (such as -1, -0.5, 0.5, 1, etc). The details for performing the tipping point analysis using the imputed datasets under the MAR assumption are provided below.

Step 1: Generate the imputed datasets under the MAR assumption (sample SAS code is provided in Section 14.3).

Step 2: For imputed values assuming MAR in the placebo group, no shift is added. For imputed values assuming MAR in the active treatment groups assumed MNAR, a shift value is added to the imputed values.

Step 3: Apply MMRM model on imputed datasets by adding the shifting value and Rubin's rule to derive the statistics and the corresponding p-values (sample SAS code is provided in Section 14.4).

14.8. Appendix H: Sample SAS Code Performing the Logistic Regression on Imputed Datasets

```
PROC LOGISTIC DATA=df_mi;
BY _Imputation_;
CLASS trt(ref='0') strvar1 strvar2/param=ref;
MODEL binary(EVENT='1') =trt Baseline strvar1 strvar2;
estimate 'TRT01PN  1 vs 0' trt 1 0 / exp cl;
estimate  'TRT01PN  2 vs 0' trt 0 1 / exp cl;
estimate  'TRT SRK-015 vs 0' trt 0.5 0.5 / exp cl;
ODS OUTPUT ESTIMATES=est_tmp;
RUN;
PROC SORT data= est_tmp; BY _IMPUTATION_; RUN;
*** Combine transformed estimates;
PROC MIANALYZE DATA= est_tmp;
ODS OUTPUT PARAMETERESTIMATES=mian_lgsodds_t;
MODELEFFECTS estimate;
STDERR stderr;
RUN;
*** Back-transform combined values;
DATA result_lgsodds;
SET mian_lgsodds_t;
Estimate_back = EXP(ESTIMATE); *Pooled odds ratio;
LCL_back=Estimate_back*EXP(-1.96*STDERR); *Pooled lower limit;
UCL_back=Estimate_back*EXP(+1.96*STDERR); *Pooled upper limit;
RUN;
```

14.9. Appendix I: Sample SAS Code Performing the CMH Test on Imputed Datasets

```
/* Obtain CMH estimate of the common odds ratio adjusted for baseline score category */;
PROC FREQ DATA=df_mi;
TABLES strvar1*strvar2*trt*binary / cmh;
ODS OUTPUT COMMONRELRISKS=comrrout CMH=cmh;
BY _Imputation_;
RUN;
/* Log-transform odds ratio estimates and obtain standard error from confidence intervals */;
DATA ormh_t;
SET comrrout(WHERE=(StudyType='Case-Control'));
log_or_mh_value=log(VALUE);
log_or_mh_se=(log(UPPERCL)-log(LOWERCL))/(2*1.96);
RUN;
/* Apply Wilson-Hilferty transformation to the CMH statistic and standardize the resulting
normal variable */;
DATA cmh_wh;
SET cmh(WHERE=(AltHypothesis='General Association'));
cmh_value_wh=((VALUE/DF)**(1/3) - (1-2/(9*DF)))/SQRT(2/(9*DF));
cmh_sterr_wh = 1.0;
RUN;
/* Combine results */;
PROC MIANALYZE DATA=cmh_wh;
ODS OUTPUT PARAMETERESTIMATES=mian_cmh_wh;
MODELEFFECTS cmh_value_wh;
STDERR cmh_sterr_wh; RUN;
/* Compute one-sided p-value */;
DATA mian_cmh_wh_p;
SET mian_cmh_wh;
IF tValue > 0 THEN Probt_upper = Probt/2;
ELSE Probt_upper = 1-Probt/2;
RUN;
```