NOVARTIS

Clinical Development

OAV101/Onasemnogene abeparvovec/Zolgensma®

COAV101A12306 / NCT04851873

A Phase IIIb, open-label, single-arm, single-dose, multicenter study to evaluate the safety, tolerability and efficacy of gene replacement therapy with intravenous OAV101 (AVXS-101) in pediatric patients with spinal muscular atrophy (SMA)

Statistical Analysis Plan (SAP)

Author:	Trial Statistician
Document type:	SAP Documentation
Document status:	Final Amendment 2
Release date:	03-Jul-2023
Number of pages:	30

Template Version 4.0, Effective from 23-Apr-2021

Document History – Changes compared to previous final version of SAP

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
8-Oct-2021	Prior to First participant treated	Creation of final version	N/A - First version	NA
18-Oct-2022	At last participant in	Protocol Amendment	Changes to protocol implemented; further details for clarification	1, 1.2, 1.2.1, 2.6.1, 2.7.1.1, 2.7.4.2, 5.6 (new), 6 following protocol amendment 1.2.1, 2.1, 2.1.1, 2.2.1, 2.3.2, 2.4.2, 2.4.2.1, 2.5.1, 2.6.2, 2.7, 2.7.3, 2.7.4.4, 2.11, 2.12, 2.13, 5.1, 5.1.2, 5.5 for clarification 2.2.1 for less subgroups 2.5.1 and 2.7.3 on laboratory parameters of primary interest 2.5.2, 2.5.5, 2.6.5, 2.7.1, 2.7.1.1, 2.7.3, 2.7.4.1, [redacted] for additional analyses
03-Jul-2023	Before DB lock	Additional and dropped analyses	Details of additional analyses; deletions	2.1.1, last visit window expanded 2.1, 2.4.2.1, 2.5.2, 2.6.2, 2.7.1, [redacted], 2.12, 5.4.2, 5.6, for revised analyses

Table of contents

List of tables

List of abbreviations

Abbreviation	Description
AAV	Adeno-Associated Virus
AAV9	Adeno-Associated Virus Serotype 9
[redacted]	[redacted]
AE	Adverse Event
AESI	Adverse Event of Special Interest
ALP	Alkaline Phosphatase
ALT	Alanine Aminotransferase
AST	Aspartate Aminotransferase
ATC	Anatomical Therapeutic Chemical
Bayley-III	Bayley Scales of Infant and Toddler Development-Third Edition
BiPAP	Bi-Level Positive Airway Pressure
BLQ	Below Limit of Quantification
CK-MB	Creatinine Kinase-MB
CMQ	Customized Medical Query
CTCAE	Common Technology Criteria for Adverse Events
DMT	Disease-modifying treatment
DNA	Deoxyribonucleic Acid
DRG	Dorsal Root Ganglia
ECG	Electrocardiogram
FAS	Full Analysis Set
HFMSE	Hammersmith Functional Motor Scale - Expanded
HIV	Human immunodeficiency virus
ICH	International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use
IV	Intravenous
kg	Kilogram(s)
MedDRA	Medical Dictionary for Drug Regulatory Activities
mg	milligram(s)
PD	Pharmacodynamic(s); Protocol Deviation
PK	Pharmacokinetic(s)
PT	Preferred Term
QTcF	QT interval corrected by Fridericia's formula
RBC	Red Blood Cell
RULM	Revised Upper Limb Module
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SAS	Statistical Analysis System
SMA	Spinal muscular atrophy
SMN	Survival of Motor Neuron
SMN1	Survival of Motor Neuron 1

SMN2	Survival of Motor Neuron 2
SMQ	Standardized Medical Query
SNAP	Sensory nerve action potential
SOC	System Organ Class
ULN	Upper Limit of Normal
vg	Vector Genome
WBC	White Blood Cell
WHO	World Health Organization
WHO-MGRS	World Health Organization Multicentre Growth Reference Study

1 Introduction

The purpose of this document is to provide further details about the statistical analysis methods, data derivations and data summaries to be employed in the study protocol COAV101A12306: A Phase IIIb, open-label, single-arm, single-dose, multicenter study to evaluate the safety, tolerability and efficacy of gene replacement therapy with intravenous OAV101 (AVXS-101) in pediatric patients with spinal muscular atrophy (SMA). This statistical analysis plan (SAP) has been based on International Conference on Harmonization (ICH) E3 and E9 guidelines and in reference to protocol version 00: dated 16 March 2021, protocol version 01: dated 28 Feb 2022, and Annotated Case Report Form: dated 25 August 2021. The statistical analysis plan covers statistical analysis, tabulations and listings of all data including efficacy and safety data.

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

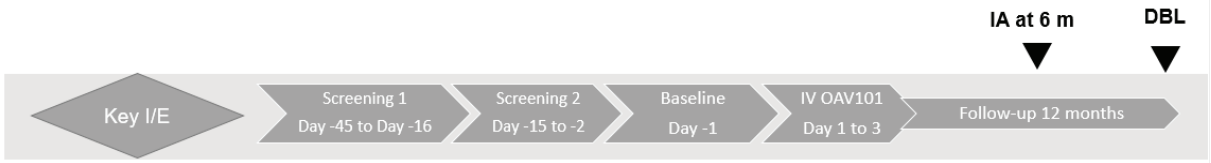
1.1 Study design

This is an open-label, single arm, multi-center study to evaluate the safety, tolerability and efficacy of intravenous (IV) OAV101 in SMA participants. The study will enroll participants that weigh ≥ 8.5 to ≤ 21 kg. Approximately 24-30 participants may be enrolled and distributed between 6 and 10 participants per cohort in each of the three brackets in order to achieve an approximately even weight distribution (≥ 8.5-13kg, > 13-17kg, > 17-21kg).

Participants who meet eligibility criteria at screening and baseline visits will receive a single dose of IV OAV101 on Day 1 (Treatment period) at the approved dose of 1.1e14 vg/kg and will be followed for a period of 12 months. The study will include a 45 day screening period in which there will be 2 screening visits, during which, eligibility will be assessed (Screening 1), weight will be collected for dose calculation (Screening 2), and baseline assessments will be performed prior to treatment. For the study duration, participants will be completing visits as defined in the Schedule of Assessments. On Day -1, participants will be admitted to the hospital for pre-treatment baseline procedures including prednisolone treatment per study protocol. On Day 1, participants will receive a single IV infusion of OAV101 and will undergo in-patient safety monitoring over the next 48 hours. Participants may be discharged 48 hours after the infusion, based on Investigator judgment.

Safety monitoring will be performed on an ongoing basis per protocol requirement and will be evaluated by the clinical safety team as well as project-level Data Monitoring Committee. Follow-up visits are planned after 1, 2, 3, 4, 6, 8, 10, 13, 26, 39, and 52 weeks.

An interim analysis for safety and efficacy is planned once the last participant completes 6-months of follow-up, and will include all available data up until that data cut-off. Final analysis will be planned after 12 months visits (end of study).



After study completion, eligible participants will be invited to enroll into Long Term follow-up study to collect additional safety and efficacy data.

1.2 Study objectives, endpoints and estimands

Objective(s)	Endpoint(s)
Primary objective(s)	**Endpoint(s) for primary objective(s)**
• The primary objective of this study is to assess the safety and tolerability of IV OAV101 over a 12-month period in participants with SMA weighing ≥8.5 kg to ≤ 21 kg.	• Evaluation of treatment emergent AEs and SAEs • Evaluation of important identified and important potential risks • Evaluate changes from baseline in vital signs, cardiac safety assessments, and clinical laboratory results
Secondary objective(s)	**Endpoint(s) for secondary objective(s)**
• Determine the efficacy of IV OAV101 at 6 and 12 months post dose in participants with SMA weighing ≥ 8.5 kg to ≤ 21 kg, as measured by secondary endpoints.	• Achievement of development motor milestones according to the World Health Organization-Multicentre Growth Reference Study (WHO-MGRS) and Bayley Scale of Infant and Toddler Development - Third Ed (Bayley-III) criteria o holds head erect for at least 3 seconds without support o sits with slight support 30 seconds o sitting without support o sits without support 30 seconds o hands-and-knees crawling o pulls to stand o standing with assistance o walking with assistance o standing alone o walking alone • Change from baseline in Hammersmith Functional Motor Scale - Expanded (HFMSE), as appropriate according to participant age. • Change from baseline in Revised Upper Limb Module (RULM), as appropriate according to participant age.

[redacted]

1.2.1 Primary estimand(s)

The clinical question of interest: What is the safety of single administration of IV OAV101 treatment in SMA participants that weigh ≥ 8.5 to ≤ 21 kg?

The primary estimand is described by the attributes listed below.

The population of interest is participants with SMA that weigh ≥ 8.5 to ≤ 21 kg

The primary variables include:

1. Treatment emergent AEs and SAEs
2. Important identified and important potential risks
3. Change from baseline in vital signs, cardiac safety assessments, and clinical laboratory results

The treatment of interest is IV OAV101.

The remaining intercurrent events include receiving prohibited concomitant medications with an impact on safety, such as concomitant medication with the intent to treat SMA. All observed AE/SAE/AESI will be reported (treatment policy strategy). For primary variable 3, if participants take prohibited concomitant medications, all the measures taken after will be disregarded as missing data and handled by a hypothetical strategy under the assumption of missing at random.

The summary measures include incidence and severity for treatment emergent AEs and SAEs as well as important identified and important potential risks and mean for changes from baseline in vital signs and clinical laboratory results.

1.2.2 Secondary estimand(s)

The secondary clinical question of interest: What is the efficacy of single administration of IV OAV101 treatment in SMA participants that weigh ≥ 8.5 to ≤ 21 kg?

The secondary estimand is described by the attributes listed below.

The population of interest is patients with SMA that weigh ≥ 8.5 to ≤ 21 kg.

The secondary variables include:

1. Change from baseline in Hammersmith Functional Motor Scale - Expanded (HFMSE), as appropriate according to participant age.

2. Change from baseline in Revised Upper Limb Module (RULM), as appropriate according to participant age.

3. Achievement of developmental motor milestones according to the World Health Organization-Multicentre Growth Reference Study (WHO-MGRS) and Bayley Scales of Infant and Toddler Development - Third Ed (Bayley-III) criteria.

- holds head erect for at least 3 seconds without support
- sits with slight support 30 seconds
- sitting without support
- sits without support 30 seconds

- hands-and-knees crawling
- pulls to stand
- standing with assistance
- walking with assistance
- standing alone
- walking alone

The treatment of interest is IV OAV101.

The remaining intercurrent events include receiving prohibited concomitant medications with an impact on efficacy, such as SMN-targeting and investigational SMA medications. Loss to follow-up due to reasons other than death will be handled by a hypothetical strategy for the change from baseline variables and by a composite strategy for milestones.

The summary measures include mean for changes from baseline in HFMSE and RULM as well as incidence for milestone achievement.

2 Statistical methods

2.1 Data analysis general information

Novartis will be performing interim and final analyses.

Analyses will be based on this document and performed using SAS® Version 9.4 (SAS Institute, Inc., Cary, NC) or later.

Categorical data will be presented as frequencies and percentages. Unless otherwise specified below, for continuous data, mean, standard deviation, median, minimum, and maximum will be presented. For continuous data summaries, values reported as below a limit will be summarized with the value of 0.5 times the limit, values reported as above a limit will be summarized with the value of 1.5 times the limit.

Where applicable, primary and secondary endpoints will be analyzed by analysis visits, [redacted]

All individual subject data will be presented in data listings.

2.1.1 General definitions

Investigational drug is the study drug OAV101 at 1.1e14 vg/kg.

Study day is the number of days from the Day 1 +1, if after Day 1, where Day 1 is the day of single administration of investigational drug. Study day is the negative number of days to Day 1, if before Day 1.

“Baseline” refers to a non-missing measurement or evaluation made prior to initiation of OAV101 infusion. If there are multiple measurements prior to the initiation of OAV101

infusion, only the latest measurement will be considered as baseline for analysis purposes. If these multiple measurements occur at the same time or time is not available, then the average of these measurements (for continuous data) or the worst among these measurements (for categorical data) will be considered as the baseline value. This same baseline value will be used throughout the study periods.

Change from baseline is defined as post-baseline value minus baseline value.

"Last contact" refers to the last study visit. If not terminated early, it is the end-of-study visit.

Analysis visits are defined following a visit windowing approach based on study day post dosing. For by-visit endpoints, the time windows describe how data will be assigned to protocol-specified time points during follow up.

Per week, the time windows for endpoints assessed at all visits are around the nominal Study Days as follows:

Visit	Nominal day	Start day	End day
Week 1	Day 8	Day 4	Day 11
Week 2	Day 15	Day 12	Day 18
Week 3	Day 22	Day 19	Day 25
Week 4	Day 29	Day 26	Day 36
Week 6	Day 43	Day 37	Day 50
Week 8	Day 57	Day 51	Day 64
Week 10	Day 71	Day 65	Day 82
Week 13	Day 92	Day 83	Day 137
Week 26	Day 183	Day 138	Day 228
Week 39	Day 274	Day 229	Day 319
Week 52	Day 365	Day 320	Day 455

For efficacy endpoints to be assessed at Weeks 4, 13, 26, 39, and 52 only, visit windows are defined as

Visit	Nominal day	Start day	End day
Week 4	Day 29	Day 4	Day 60
Week 13	Day 92	Day 61	Day 137
Week 26	Day 183	Day 138	Day 228
Week 39	Day 274	Day 229	Day 319

Week 52	Day 365	Day 320	Day 455

For ECG and echocardiogram endpoints to be assessed at Weeks 1, 4, 8, 13, 26, and 52 only, visit windows are defined as

Visit	Nominal day	Start day	End day
Week 1	Day 8	Day 4	Day 18
Week 4	Day 29	Day 19	Day 42
Week 8	Day 57	Day 43	Day 74
Week 13	Day 92	Day 75	Day 137
Week 26	Day 183	Day 138	Day 273
Week 52	Day 365	Day 274	Day 455

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

An "Adverse Event" (AE) is any untoward medical occurrence in a clinical investigation subject, which does not necessarily have a causal relationship with the drug or device under study.

A "treatment-emergent Adverse Event" is any AE whose onset or worsening occurred on or post Day 1.

2.2 Analysis sets

The Full Analysis Set (FAS) and Safety Set both comprise all participants that received any investigational drug.

The enrolled population will consist of all participants who complete the screening period. For rules of exclusion criteria of analysis sets see Appendix 5.5.

2.2.1 Subgroup of interest

Weight brackets (≥ 8.5-13 kg, > 13-17 kg, > 17-21 kg) are subgroups of interest. Weight brackets are defined by body weight measurements at Screening visit 2. Data will be summarized by weight bracket and overall.

2.3 Patient disposition, demographics and other baseline characteristics

2.3.1 Patient disposition

Disposition will be presented for all participants, which will include the following:

- The number of participants screened
- The number (%) of screened participants who screen failed and the reasons for screen failure (inclusion/exclusion criteria, withdrew consent, and/or other) will be summarized. A listing of reasons for screen failure will be provided for all participants who screen failed.
- The number (%) of participants in each weight bracket (≥ 8.5-13 kg, > 13-17 kg, > 17-21 kg), and overall, for the FAS.
- The number (%) of participants who completed the study, defined as all participants with a final scheduled visit at Month 12 conducted either remotely or at the study site
- The number (%) of participants who discontinued from the study and the associated reasons

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

2.3.2 Demographics and other baseline characteristics

The age of the participant at the time of OAV101 infusion will be summarized. The distribution of participants by sex, race, and ethnicity will be presented. If a participant reports multiple races, they will be categorized as "multiple races" in the summary table. Participant demographics will be summarized by weight bracket (≥ 8.5-13 kg, > 13-17 kg, > 17-21 kg) and overall, for the FAS.

Demographic data will be determined using the following calculations:

If age at Screening visit 1 will be given in months or years, respectively, age in days is 30.4375 times age in months and 365.25 times age in years.

Age at Study day 1 = (Study day 1 visit date – Screening visit 1 date + age at Screening visit 1 in days)/365.25 days, expressed in years

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

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

The following statistics regarding the participant's characteristics at birth will be summarized: gestational age (weeks).

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

The time from SMA diagnosis to study start at Screening visit 1 is defined as the difference in days between the dates. The same approach is taken for time since first symptoms accordingly.

Age at SMA diagnosis in years is calculated as age at study start in days minus time from SMA diagnosis to study start, divided by 365.25. If diagnosis is prenatal, age at SMA diagnosis is 0 years. If diagnosis is not prenatal or unknown to be prenatal, minimum age at SMA diagnosis is 0 years.

Age at first symptoms of SMA in years is calculated as age at study start in days minus time from first symptoms of SMA to study start, divided by 365.25. Minimum age at first symptoms of SMA is 0 years.

Where appropriate, frequencies of age groups will be displayed as 0-<28 days, 28 days-<2 years, 2-<12 years, 12-<18 years, 18 years or more.

Baseline characteristics including virus SMN1 details including presence/absence of bi-allelic deletion, point mutations, and genetic modifier G>C.859, SMA type, SMN2 copy number, pre-treatment with SMA disease-modifying treatments (DMT), developmental motor milestone status, and total scores of HFMSE and RULM will be summarized by weight bracket (≥ 8.5-13 kg, > 13-17 kg, > 17-21 kg) and overall, for the FAS. Virus serology variables include HIV, hepatitis B and hepatitis C test results. The coagulation panel includes prothrombin time (PT), International Normalized Ratio (INR) and activated partial thromboplastin time (aPTT). Duration of SMA DMT will be summarized.

Relevant medical histories and current medical conditions at baseline will be summarized by system organ class and preferred term by weight bracket and overall.

2.4 Treatments (study treatment, rescue medication, concomitant therapies, compliance)

2.4.1 Study treatment / compliance

The Safety set will be used for the analyses below. Categorical data will be summarized as frequencies and percentages. For continuous data, descriptive statistics including mean, standard deviation, median, minimum, and maximum will be presented.

The duration of administration in minutes, volume of infusion, dose administration pause during infusion, and treatment compliance (%) will be summarized by means of descriptive statistics using the safety set. Treatment compliance is defined as 100 times the dose administered divided by the planned dose, 1.1e14 vg/kg body weight.

2.4.2 Prior, concomitant and post therapies

Prior and concomitant medications and significant non-drug therapies prior to and after the start of the study treatment will be listed and summarized according to the ATC classification system.

A prior medication is defined as any medication taken prior to the date of the OAV101 infusion. A concomitant medication is defined as medication that started prior to the date of the OAV101 infusion and continued to be taken after the infusion or any medication that started on or after the date of the OAV 101 infusion. The number and percentage of participants in the safety analysis set taking prior or concomitant medications will be summarized by generic drug name

based on the WHO Drug Dictionary. Further summaries will present prior medications ongoing at Study Day 1.

2.4.3 Specific Medication Subgroups

To reduce the host immune response to the AAV-based therapy, participants will receive prophylactic prednisolone or equivalent steroid (approximately 1 mg/kg/day) 24 hours prior to the gene transfer and continuing for approximately 30 days. After 30 days of treatment, the dose of prednisolone or equivalent can be tapered for participants whose ALT values and AST values are ≤2 X ULN.

The total number of days receiving prophylactic prednisolone or equivalent and total cumulative dose of prophylactic prednisolone administered during the study (mg/kg, derived from the prednisolone log) will be computed for each participant. Partial treatment start and end dates will be imputed as described in section 5.1.3. In case when other glucocorticosteroids were used, the doses should be converted into Prednisolone equivalent doses according to the relative potency, for example with a factor of 0.25 for hydrocortisone, 6.67 for dexamethasone and 1.25 for methylprednisolone. Dose levels at end of study will be summarized, frequencies of participants with >1mg/kg at any time, >2mg/kg at any time, with dose interruptions of more than 7 days and with any intravenous dosing will be reported. The number and percentage of participants with versus without prednisolone equivalent dose will also be presented per visit and weight bracket.

To compute total cumulative dose, the total dosing period is subdivided into dosing intervals represented by constant dose levels. Total doses per dose level and dosing interval are derived as the product of dose level in mg/kg and the number of days of the dosing interval. For the total cumulative dose, all products per dose level and dosing interval are summed up.

Exposure will be summarized for the Safety Analysis Set.

2.5 Analysis supporting primary objective(s)

2.5.1 Primary endpoint(s)

The primary endpoints include:

1. Incidence and severity of treatment emergent AEs and SAEs
2. Incidence of important identified and important potential risks
3. Change from baseline in vital signs, clinical laboratory, and procedure (eg ECG, echocardiogram) results

An AE is treatment-emergent, if it occurs or worsens after the start of the infusion of the investigational drug.

Important identified and important potential risks include the following AEs (see section 2.7.1.1 for detail):

- Hepatotoxicity
- Thrombocytopenia
- Cardiac adverse events
- Dorsal root ganglia toxicity

- Thrombotic microangiopathy

Vital sign parameters include height, weight, body mass index, blood pressure, respiratory rate, pulse, temperature, and oxygen saturation level. On Day 1, vital signs will be recorded pre-dose and then monitored post dose and at each visit.

Laboratory parameters in the scope of primary endpoints include platelets, troponin-I, and among chemistry variables alanine aminotransferase (ALT), aspartate aminotransferase (AST), glutamate dehydrogenase (GLDH), alkaline phosphatase (ALP), gamma glutamyl transferase (GGT), serum total bilirubin, direct bilirubin, and indirect bilirubin.

Cardiac procedures include ECG and echocardiogram. ECG parameters include QT, RR, and QTcF. Reviewed by local and central cardiologists, echocardiogram parameters include thrombus status, left ventricular ejection fraction (%), left ventricular fractional shortening (%), and left ventricular global longitudinal strain. ECG overall interpretation will be given.

2.5.2 Statistical hypothesis, model, and method of analysis

No hypothesis testing will be performed.

Incidence and severity for treatment emergent AEs and SAEs as well as important identified and important potential risks will be summarized by weight bracket and overall in the safety set. The number and proportion of participants reporting a treatment emergent AE or SAE at least once will be reported. Summaries will also be provided by MedDRA System Organ Class and Preferred Term.

Changes from baseline in vital signs and clinical laboratory results will be summarized descriptively by weight bracket and overall in the safety set. Mean, standard deviation, median, minimum, and maximum will be presented. For each applicable time point, only participants with a baseline value and a measurement for that time point will be included in the summary. Shift tables of laboratory abnormality versus baseline will be provided for minimum, maximum and final post-baseline values. For laboratory parameters of primary interest except for GLDH, direct and indirect bilirubin, additional shift tables will consider CTCAE grading, version 5. For troponin-I, any value above the upper limit of the reference range (ULN) is considered CTCAE Grade 1. Number of participants with ALT >ULN, 2xULN, 3xULN per visit will be summarized. Individual courses of and median changes from baseline in clinical laboratory results of ALT, AST, GLDH, ALP, GGT, serum total bilirubin, direct bilirubin, indirect bilirubin, platelets, and troponin-I over time will be visualized. Graphs of median changes will be per weight bracket. Graphs of individual clinical laboratory results will use normalized values as multiples of ULN.

Procedure results including ECG QTcF and overall interpretation will be summarized descriptively by weight bracket and overall.

2.5.3 Handling of intercurrent events

For AEs, SAEs, important identified and important potential risks, intercurrent events will not be considered following a treatment policy strategy.

Use of prohibited concomitant medications with an impact on safety: Data collected while/after receiving prohibited concomitant medications with an impact on safety will be included in the analyses of AEs (treatment policy strategy).

Study discontinuation due to reasons other than death: Data collected up to the point of discontinuation will be included in all primary endpoint analyses.

Study discontinuation due to death: It is anticipated that this intercurrent event is unlikely to occur during the study. Data collected up to the point of death will be included in all primary endpoint analyses.

Changes from baseline in vital signs, clinical laboratory, and procedure results will be summarized for all data available, only excluding data collected after the start of intake of prohibited medication, if any, following a hypothetical strategy and assuming this intercurrent event occurring at random.

2.5.4 Handling of missing values not related to intercurrent event

Missing values not related to intercurrent events will not be imputed. All data of participants in the safety set will be presented as observed, collected and derived.

2.5.5 Sensitivity analyses

Given the descriptive nature of the summary of primary safety endpoints, no sensitivity analysis is planned. Only if intercurrent events of intake of prohibited concomitant medication with the intent to treat SMA will occur, in a sensitivity analysis of changes from baseline in vital signs, ECG and laboratory parameters all data including those collected after the intercurrent event started will be considered following a treatment policy strategy.

2.5.6 Supplementary analyses

No supplementary analysis is planned.

2.6 Analysis supporting secondary objectives

2.6.1 Secondary endpoint(s)

The secondary endpoints include:

1. Changes from baseline in HFMSE and RULM;

2. Percent of participants achieving development motor milestones according to the WHO-Multicenter Growth Reference Study (WHO-MGRS) and Bayley Scale of Infant and Toddler Development - Third Ed (Bayley-III) criteria.

For all participants of 24 months of age or more at the time of visit, HFMSE (2019 edition) will be assessed, consisting of 33 SMA-specific items with a maximal total score of 66 (O'Hagen et al 2007). Each item is scored on a 3-point Likert scale from 0 (no response) to 2 (full response) and summed up to the total score. A higher score indicates a higher ability level. A total score will not be derived if any item score is missing or "cannot test".

For all participants of 30 months of age or more at the time of visit, RULM will be assessed, consisting of 19 items with a maximal total score of 37 with lower scores reflecting poorer ability. RULM is an assessment that measures motor performance in the upper limbs from childhood to adulthood in ambulatory and non-ambulatory individuals with SMA. There are 18 items scored 0 (unable) to 2 (full achievement) and one item scored 0 (unable) to 1 (able) summed up to the total score (Pera et al 2019). For participants of 24 to 30 months of age at the time of visit, the two test items (C)- Tracing a path and (H)- Tearing Paper will not be administered and therefore for a participant between the ages of 24-30 months the maximum score achieved is 33 points. A total score will not be derived if any item score is missing or "cannot test".

Changes from baseline are only summarized for subjects whose baseline values are available. For younger subjects who are not eligible for RULM/HFMSE measure at baseline, their scores at later visits will only be listed.

The Developmental Milestone Checklist consists of 10 Developmental Motor Milestones obtained from WHO-MGRS (sitting without support, hands-and-knees crawling, standing with assistance, walking with assistance, standing alone, and walking alone) and from Bayley-III (head control (Bayley gross motor item #4), sits with slight support for at least 30 seconds (Bayley gross motor item #19), sits without support for at least 30 seconds (Bayley gross motor item #26), and pulls to stand (Bayley gross motor item #35)) (WHO Multicentre Growth Reference Study Group 2006, Bayley 2006).

2.6.2 Statistical hypothesis, model, and method of analysis

No hypothesis testing will be performed.

The secondary endpoints will be summarized by weight bracket and overall for the FAS.

For all participants of 24 (HFMSE) or 30 (RULM) months of age or more at baseline, changes from baseline in HFMSE and RULM will be summarized descriptively; in addition, mixed models with repeated measurement will be fitted that include the change from baseline as the dependent variable, and fixed effects of weight bracket, visit, and their interaction as well as covariate of baseline value. The least squares means and 95% 2-sided confidence intervals will be reported for each visit. Changes in total scores over time will be visualized by means of box-whisker plots and individual longitudinal trajectories. The number and proportion of participants with increases and with decreases in total scores compared to baseline (any, by more than 3, 6, 10 points inclusive) will be presented by visit and weight bracket.

The number and proportion of participants achieving each developmental motor milestone will be presented per visit by weight bracket and overall. In addition, summaries of achievements respective to baseline achievement status will also be generated separately.

2.6.3 Handling of intercurrent events

Changes in HFMSE and RULM and developmental motor milestone achievements will be summarized for all data available, only excluding data collected after the start of intake of prohibited medication with an impact on efficacy, if any, following a hypothetical strategy and assuming this intercurrent event occurring at random.

Following a composite strategy, in participants of the FAS dying before the achievement of a developmental motor milestone, the milestone is assumed not being achieved at any of the visits planned after the death occurred.

2.6.4 Handling of missing values not related to intercurrent event

Missing values not related to intercurrent events will not be imputed. All data of participants in the FAS will be presented as observed, collected and derived.

2.6.5 Sensitivity analyses

For HFMSE and RULM, in a sensitivity analysis of change in total score from baseline, on any missing data due to SMA-related death or invasive ventilation the HFMSE and RULM total scores will be set to 0, and on any missing data due to death for other causes to the last observed total score before death.

In further sensitivity analyses of changes from baseline in HFMSE and RULM total scores, missing item scores and "cannot test" assessments will be imputed by 0 scores for analysis. This applies as long as no more than 6 (HFMSE) and 3 (RULM) item scores per visit are missing or "cannot test". If more items are missing or "cannot test", total scores are not derived.

2.6.6 Supplementary analyses

No supplementary analysis is planned.

2.7 Safety analyses

Safety will be assessed through the incidence and severity of AEs, vital sign assessments, cardiac assessments, laboratory evaluations (chemistry, hematology, immunology, urinalysis), physical and neurological examinations, and use of concomitant medications. Adverse events will be coded in accordance with the most current version of the MedDRA coding dictionary. Safety assessments will be presented in data listings or patient profiles for the safety population.

Safety is the primary objective of the study. Summary of adverse events, vital signs, and clinical laboratory results are described in Section 2.5. Summaries will be presented per weight bracket and overall.

2.7.1 Adverse events (AEs)

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

Summary tables by SOC and PT will be provided for

- Number (%) of participants with treatment-emergent AEs
- Number (%) of participants with treatment-emergent AEs by severity
- Number (%) of participants with treatment-emergent SAEs
- Number (%) of participants with treatment-emergent AEs leading to discontinuation
- Number (%) of participants with treatment-emergent AEs leading to death
- Number (%) of participants with drug-related AEs
- Number (%) of participants with drug-related SAEs
- Number (%) of participants with non-serious treatment-emergent AEs.

For all adverse event summaries, the number of treatment-emergent adverse events, the number and percentage of participants experiencing treatment-emergent adverse events and the time-adjusted rate of occurrence will be tabulated according to SOC and PT. Participants reporting more than one adverse event for a given PT will be counted only once for that term (most severe incident for the severity tables and most related incident for the relationship tables). Participants reporting more than one adverse event within a SOC will be counted only once for that SOC. Participants reporting more than one adverse event will be counted only once in the overall total. Rates of occurrence are adjusted by the sum of follow-up time across all participants, which is the time at risk from dosing to the start of first occurrence of event in those with event and to the end of follow-up, or data cutoff date, in those without the event.

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

- Any treatment-emergent adverse event;
- Any treatment-emergent adverse event “related” to OAV101;
- Any severe treatment-emergent adverse events;
- Any serious treatment-emergent adverse events;
- Serious treatment-emergent adverse events “related” to OAV101;
- Treatment-emergent adverse events leading to discontinuation of participant from study;
- Treatment-emergent adverse events leading to death;
- Treatment-emergent adverse events of special interest (cp. Section 2.7.1.1).

AEs and SAEs are primary endpoints. For details, see Section 2.5.

2.7.1.1.1 Adverse events of special interest / grouping of AEs

On the basis of important identified and important potential risks associated with OAV101, AEs of special interest (AESI) are identified by using Standard MedDRA queries (SMQ) with terms as follows:

Hepatotoxicity as hepatic disorders (SMQ), thrombocytopenia as transient thrombocytopenia (CMQ), dorsal root ganglia toxicity as DRG cell inflammation (CMQ), and cardiac events as ischemic heart disease (SMQ) or cardiomyopathy (SMQ) or cardiac arrhythmias (SMQ) or embolic and thrombotic events (SMQ) or myocardial infarction (SMQ), respectively.

Thrombotic microangiopathy is identified via the following approach:

- Criteria #1: cases with any one of the following PTs: thrombotic microangiopathy OR haemolytic uraemic syndrome OR atypical haemolytic uraemic syndrome.
- Criteria #2: cases with at least one PT from EACH of the following SMQs representing thrombocytopenia, hemolysis and relevant renal events respectively:
    - Haematopoietic thrombocytopenia (SMQ)
    - Haemolytic disorders (SMQ)
    - Acute renal failure (SMQ) OR Renovascular disorders (SMQ)

Summary tables for AESIs, will include severe AESIs, serious AESIs, and drug-related AESIs. All AESIs will be listed together with a summary of prednisolone equivalent dose regimens over time with study days of start and end of AESI and per dose of prednisolone equivalent indicated. The total number of days with prednisolone equivalent dosing will be derived.

2.7.2 Deaths

Deaths will be presented in data listings and summary tables as AEs with fatal outcome. All data will be considered for analysis up to the time of death.

See Section 2.6 on how composite strategy is used to consider death for development milestone achievements.

2.7.3 Laboratory data

Safety laboratory data generated by the central laboratory will be used in all analyses. Safety laboratory data generated by local lab will only be listed. In data listings, per parameter, values will also be given as multiples of upper normal range limits.

Laboratory parameters include hematology, chemistry and urinalysis variables.

Changes in selected safety laboratory variables are primary endpoints. For details, see Section 2.5.

Hematology variables include: hematocrit, hemoglobin, red blood cell (RBC) count, white blood cell (WBC) count, differential blood count (in measured values and % of WBC; neutrophils, lymphocytes, monocytes, eosinophils, basophils, bands), mean corpuscular volume

(MCV), mean corpuscular hemoglobin (MCH), mean cell hemoglobin concentration (MCHC), and platelets.

Chemistry variables include: ALT, ALP, AST, blood urea nitrogen (BUN), calcium, creatinine, gamma glutamyl transferase, glucose, serum total bilirubin, direct bilirubin, indirect bilirubin, total creatine kinase (CK), chloride, phosphorus, sodium, potassium, lactate dehydrogenase (LDH), GLDH, troponin-I, and bicarbonate.

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

Clinical laboratory test results will be summarized at each visit including End of Study/Early Termination, for minimum, maximum and final post-baseline values. Urinalysis variables will only be listed.

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

The number and percentage of participants meeting the criteria for ALT and bilirubin defined in Table 16-1 of the protocol will be summarized by weight bracket and overall including potential Hy's Law cases. Potential Hy's Law cases will be presented for maximum post-baseline values as combined, even if criteria are fulfilled at different times during the follow-up.

A participant or event will be counted if the post-baseline laboratory values meet the above criteria regardless of the baseline laboratory value (i.e., the post-baseline laboratory value does not need to be worse than the baseline laboratory value). For participants meeting any elevation criterion, a corresponding listing of all ALT, AST, alkaline phosphatase, INR, GLDH, and total, direct, and indirect bilirubin values will be provided.

2.7.4 Other safety data

2.7.4.1.1 ECG and cardiac imaging data

Changes in cardiac safety data such as ECG and echocardiogram are primary endpoints. For details, see Section 2.5.

ECG and echocardiogram results will be summarized descriptively by weight bracket and overall. Summaries according to QTcF notable ranges (QTcF > 500 ms) will be provided.

2.7.4.1.2 Vital signs

Changes in vital signs are primary endpoints. For details, see Section 2.5. Summaries according to abnormality assessments will be provided. Vital sign results will be flagged as clinically significant if they meet the pre-specified criteria which are defined in Appendix 5.6.

Height and weight will be measured (supine or standing) at each study visit. Screening weight shall be obtained ≤ 14 days prior to Day 1. On Day 1, weight and height will be measured pre-dose. Changes in height, weight and body mass index will be presented in summary tables over time.

2.7.4.1.3 Neurological examination and sural sensory nerve action potential

Abnormal findings resulting from the examination of proprioceptive, vibratory, tactile and pain sensation will be summarized.

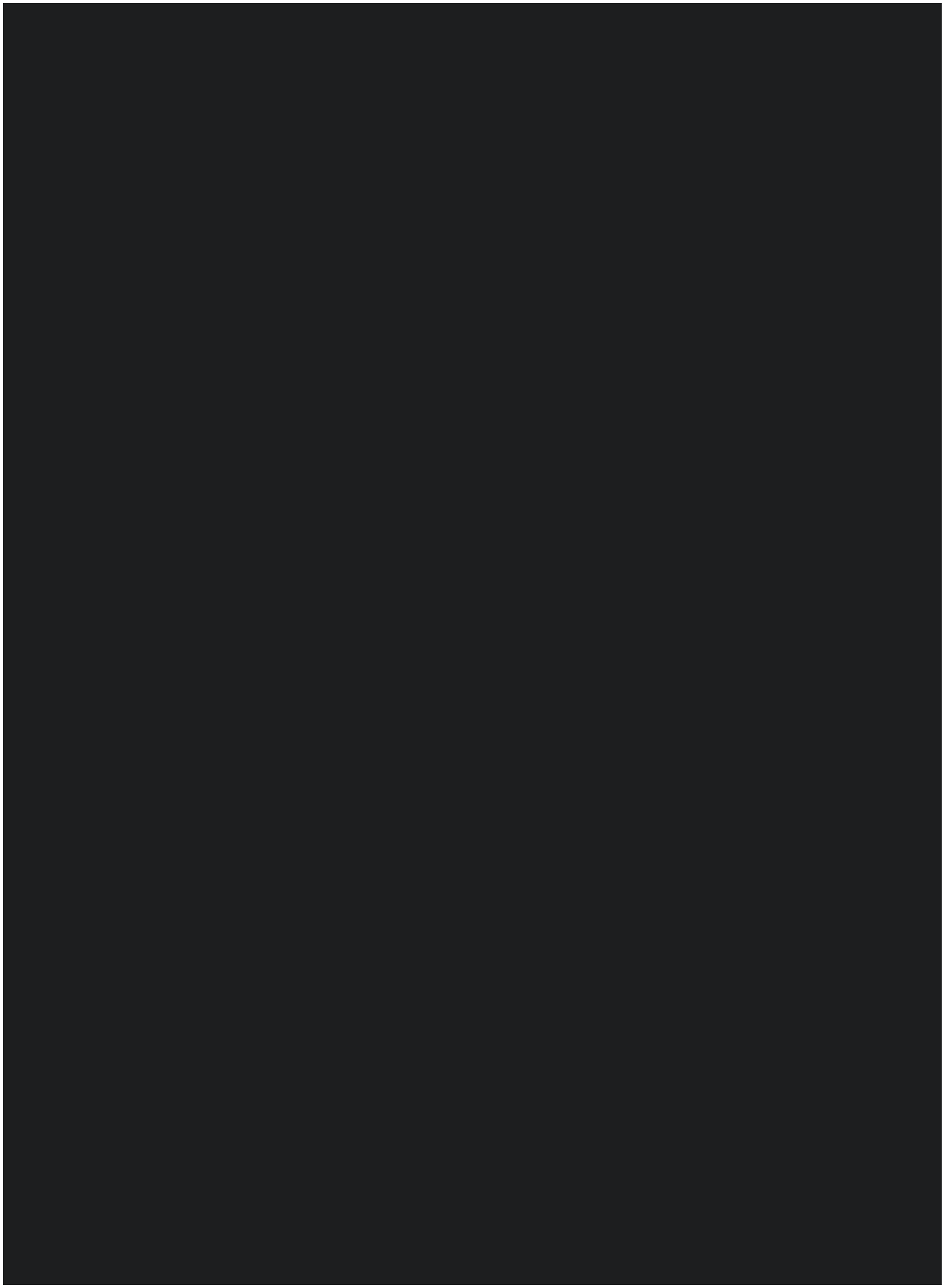
To complement neurological examination in case of clinically significant sensory abnormalities, sural sensory nerve action (SNAP) potential will be assessed. SNAP parameters include sural, median and peroneal SNAP voltage and velocity as well as overall interpretation, presented in data listings.

2.8 Pharmacokinetic endpoints

Not applicable.

2.9 PD and PK/PD analyses

Not applicable.



2.13 Interim analysis

An interim analysis will be performed once the last participant reaches the 6-month time point and will include all available data up until that data cut-off. The interim analysis is intended to provide an early safety readout from the study. No study design change or adaptation will be implemented based on the outcome of the interim analysis.

The interim analysis will cover all primary and secondary endpoints.

3 Sample size calculation

The size of approximately 24 participants is considered reasonable to provide descriptive safety information across 3 weight brackets ≥ 8.5 kg and ≤ 21 kg. A sample size of approximately 24 participants will provide 90% probability to observe at least one event if the underlying incidence of the event is 9%.

4 Change to protocol specified analyses

No change from protocol specified analysis was made.

5 Appendix

5.1 Imputation rules

Missing values for safety endpoints, including but not limited to AE, laboratory values, will not be imputed. For local lab, missing units or normal ranges will be reviewed by the medical team.

5.1.1 Study drug

If the date of study drug application will be missing or partial, it will be imputed by the date of the Day 1 visit.

Start time and duration of infusion of investigation drug will not be imputed if missing.

5.1.2 Date imputation

Missing or incomplete AE start and end dates will be imputed according to the general Novartis imputation rules.

Missing or incomplete dates of SMA diagnosis and of first symptoms will be imputed according to the general Novartis imputation rules.

Missing or incomplete start and end dates of medications will be imputed according to the general Novartis imputation rules.

Other data will not be imputed.

5.2 AEs coding/grading

Verbatim terms of AEs, including important identified and important potential risks, will be encoded by means of MedDRA according to the data management plan.

5.3 Laboratory parameters derivations

Hy's law criteria apply (see Section 2.7.3).

5.4 Statistical models

5.4.1 Analysis supporting primary objective(s)

Not applicable.

5.4.2 Analysis supporting secondary objective(s)

A mixed model of repeated measures is applied using unstructured covariance matrix as long as the model converges (compare Section 2.6.2). If the model does not converge with unstructured covariance matrix, heterogeneous compound symmetry of the covariance matrix will be assumed.

5.5 Rule of exclusion criteria of analysis sets

Table 1 Criteria leading to exclusion

Analysis Set	Criteria that cause subjects to be excluded
Enrolled Set	Not having informed consent; Screening failed
FAS, Safety Set	Not in Enrolled Set; No investigational drug taken

5.6 Clinically significant vital sign values

Vital sign results will be flagged as clinically significant if they meet the pre-specified criteria which are defined as follows.

Systolic and diastolic blood pressure:

- For participants aged <18 years at the time of assessment, systolic and diastolic blood pressure values ≥ the corresponding to the 90th percentile value for age, gender, and height will be used to classify "high" clinically significant values according to Flynn 2017, Tables 4 for boys and 5 for girls, respectively.
  - Participants aged ≥2 to <3 years at the time of assessment will utilize the blood pressure levels defined for age 2 years, participants aged ≥3 to <4 years at the time of assessment will utilize the blood pressure levels defined for age 3 years, etc. For participants <2 years of age at the time of assessment levels defined for age 1 year apply.
  - Height percentile will be defined using the height at the time of assessment, or the most recently recorded height prior to the assessment if the height at the time of assessment is not available. Participants with a height which falls between two percentiles will be classified according to the lower of the two percentiles. Participants with height below the 5th percentile will be classified according to the 5th percentile.
- For all participants regardless of age at assessment, systolic blood pressure values below the 5th percentile will be classified as "low" clinically significant values. The 5th percentile value can be estimated as 70 mmHg plus twice participant's age in years.

Temperature:

- For participants <18 years at the time of assessment, temperature values ≥38.4°C will be classified as "high" clinically significant.
- For all participants regardless of age at assessment, temperature values ≤35°C will be classified as "low" clinically significant.

Pulse rate:

The criteria for classifying high and low pulse rate values are contained in the table below. (Flemming 2011).

Age at assessment	High (bpm)	Low (bpm)
2 to <3 years	>128	<92
3 to <4 years	>123	<86
4 to <6 years	>117	<81
6 to <8 years	>111	<74
8 to <12 years	>103	<67
12 to <15 years	>96	<62
15 to <18 years	>92	<58

Weight:

- For participants <18 years at the time of assessment, weight values which reflect an increase from baseline of ≥2 BMI-for-age percentile categories (significant weight gain) will be classified as "high" clinically significant; weight values which reflect a decrease from baseline of ≥2 BMI-for-age percentile categories (significant weight loss) will be classified as "low" clinically significant.
    - Baseline BMI-for-age weight status categories are underweight (less than the 5th percentile), healthy weight (5th percentile to less than the 85th percentile), overweight (85th to less than the 95th percentile) and obese (equal to or greater than the 95th percentile).
    - BMI-for-age percentiles are obtained from the WHO Growth Charts (https://www.who.int/toolkits/child-growth-standards/standards/body-mass-index-for-age-bmi-for-age for 0 to under 5 years of age and https://www.who.int/tools/growth-reference-data-for-5to19-years/indicators/bmi-for-age for 5 to 19 years of age, respectively).

Respiratory rate:

The criteria for classifying high and low respiratory rate values are contained in the table below. (Flemming 2011; Eldridge 2014; Kou).

Age at assessment	High (breath/min)	Low (breath/min)
2 to <3 years	>34	<22
3 to <4 years	>29	<21
4 to <6 years	>27	<20
6 to <8 years	>24	<18
8 to <12 years	>22	<16
12 to <15 years	>21	<15
15 to <18 years	>20	<13

6 References

Bayley N: The Bayley-III is a revision of the Bayley Scales of Infant Development – Third Edition (BSID-III), Technical Manual. San Antonio, TX: Harcourt, 2006.

Eldridge L: What is a Normal Respiratory Rate? Updated May 16,2014.

Flemming S, Thompson M, Stevens R, et al. Normal ranges of heart rate and respiratory rate in children from birth to 18 years of age: a systematic review of observational studies. Lancet 2011; published online March 15. DOI: 10.1016/S0140-6736(10)62226-X.

Flynn JT, Kaelber DC, Baker-Smith CM, et al. Clinical Practice Guideline for Screening and Management of High Blood Pressure in Children and Adolescents. Pediatrics 2017; 140(3):e2017 1904.

ICH E9 (R1) Harmonized Guideline: addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. Final version on 20 November 2019.

Kou R, Shuei L, Bradypnea, Department of Psychology, School of Medicine, National Yang-Ming University, Taipei, Taiwan. http://rd.springer.com/referenceworkentry/10.1007%2F978-3-540-29676-8_246.

O'Hagen JM, Glanzman AM, McDermott MP, Ryan PA, Flickinger J, Quigley J, Riley S, Sanborn E, Irvine C, Martens WB, Annis C, Tawil R, Oskoui M, Darras BT, Finkel RS, De Vivo DC: An expanded version of the Hammersmith Functional Motor Scale for SMA II and III patients. Neuromuscul Disord. 2007 Oct;17(9-10):693-7. Epub 2007 Jul 19.

Pera MC, Coratti G, Mazzone ES, Montes J, Scoto M, De Sanctis R, Main M, Mayhew A, Lofra RM, Young SD, Glanzman AM, Duong T, Pasternak A, Ramsey D, Darras B, Day JW, Finkel RS, De Vivo DC, Sormani, MP, Bovis F, Straub V, Muntoni F, Pane M, Mercuri E, and the ISMAC Consortium Group: Revised upper limb module for spinal muscular atrophy: 12 month changes. Muscle Nerve 59: 426-430, 2019.

WHO Multicentre Growth Reference Study Group: WHO Motor Development Study: Windows of achievement for six gross motor development milestones. Acta Paediatrica Suppl 450: 86-95, 2006.