



## Statistical Analysis Plan

**Title: A Phase 3, Randomised, Placebo-Controlled Trial of Arimoclomol in Amyotrophic Lateral Sclerosis**

Protocol Number: ORARIALS-01  
ClinicalTrials.gov Identifier: NCT03491462

Version 2.0

Issue Date: *19-APR-2021*

Author: [REDACTED], Ph.D,  
[REDACTED], MS

Previous Versions

*Version 1.0, Issue Date 23-FEB-2021*

**Clinicaltrials.gov ID: NCT03491462**

**This document contains information which is the property of Zevra Denmark A/S and is provided here as part of the results registration on clinicaltrials.gov. It is understood that this information cannot and will not be disclosed to others without written approval from Zevra Denmark A/S.**



## SAP Amendments before database lock

Version	Issue Date	Section	Revision/Addition	Rationale
1.0	19-APR-2021	8.4 & 8.5.2	Added additional sensitivity analysis to CAFs and ALSFRS-R	During the review of blinded data of ALSFRS-R prior to unblinding, it was discovered that quite a few patients had assessments of ALSFRS-R at Visit 22 (safety follow-up) after having had a scheduled Visit 21 (Landmark at week 76) with an assessment of the ALSFRS-R. According to protocol (version 7, 08-Jun-2020) ALSFRS-R should not be assessed at the safety follow-up and the SAP's definition of the end-of the in-trial period, for reasons discussed in detail herein, ends at the Landmark Visit 21. To assess the impact of including Visit 22 ALSFRS-R assessments that were excluded from primary analyses due to in-trial period definition. Therefore, two additional post hoc sensitivity analyses are specified – one for ALSFRS-R per se and one for CAFS.
1.0		6.2	Include details for windowing weight assessments and clarify windowing for Week 76.	Current windowing does not provide details for windowing of weight assessments and details for Week 76 windowing may be confusing.



1.0		8.12.3	Corrected the PCS High the vital sign parameters PCS table (Pulse, DBP, SBP) and removed Temperature.	SAP 1.0 incorrectly displays “< X” value when should display “>X” for Pulse, DBP, SBP and temperature should have been removed prior signing version 1.
1.0		Appendix 13	Update TOC	Additions/corrections were made.



## Table of Contents

1	INTRODUCTION .....	7
2	STUDY OBJECTIVES .....	7
2.1	Primary objective .....	7
2.2	Secondary objectives .....	7
2.3	Exploratory Objectives .....	7
2.4	Safety Objective.....	7
3	ENDPOINTS .....	8
3.1	Primary Endpoint.....	8
3.2	Secondary Endpoints .....	8
3.3	Exploratory Endpoints .....	8
3.4	Safety Endpoints .....	8
4	ANALYSIS POPULATIONS, OBSERVATION PERIODS AND CENSORING .....	9
4.1	Analysis populations.....	9
4.2	Observation Periods .....	9
4.2.1	In-Trial .....	10
4.2.2	On-Treatment.....	12
4.2.2.1	Safety .....	12
4.2.2.2	Efficacy .....	12
4.3	Observation periods and population sets .....	13
5	ESTIMANDS .....	13
5.1	Primary Estimand.....	13
5.2	Secondary Estimand.....	14
5.3	While on treatment/alive estimands.....	14
6	VISIT WINDOWING AND BASELINE .....	16
6.1	Baseline.....	16
6.2	Windowing.....	16
7	SAMPLE SIZE .....	18
8	STATISTICAL ANALYSIS AND REPORTING.....	19
8.1	Introduction.....	19
8.1.1	Decimal Places.....	20
8.2	Statistical models .....	20
8.3	Primary endpoint and analyses .....	21



8.3.1	ALS Functioning Rating Scale – Revised (ALSFRS-R) .....	21
8.3.2	Time to PAV/Tracheostomy/Death .....	22
8.3.3	CAFS: Combined Assessment of Function and Survival.....	22
8.3.3.1	Generic scoring algorithm.....	22
8.3.3.2	Primary endpoint.....	23
8.3.3.3	Analysis.....	23
8.4	Sensitivity analyses of the primary endpoint .....	23
8.4.1	Subgroup analyses .....	26
8.5	Secondary endpoints and analyses.....	26
8.5.1	Time to event .....	26
8.5.2	ALSFRS-R.....	27
8.5.3	% predicted SVC.....	27
8.5.4	Subgroup analyses .....	29
8.6	Other efficacy endpoints .....	29
8.6.1.1	Edinburgh Cognitive and Behavioral ALS Screen (ECAS) .....	29
8.6.1.2	Schedule for the Evaluation of Individual Quality of Life-Direct Weighting (SEIQoL-DW).....	30
8.6.1.3	EQ-5D-5L .....	30
8.7	Multiplicity .....	31
8.8	Columbia-Suicide Severity Rating Scale.....	31
8.9	Other reporting issues .....	32
8.9.1	Derived Data .....	32
8.9.2	Patient Disposition .....	36
8.9.3	Protocol Deviations.....	36
8.9.4	Baseline Comparability .....	37
8.9.5	Medical History .....	37
8.9.6	Prior and Concomitant Medications .....	37
8.9.7	Exposure to Study Drug.....	37
8.9.8	Treatment Compliance .....	38
8.10	Pharmacokinetic Analyses .....	38
8.11	Biomarker Analyses.....	38
8.12	Safety Analyses.....	38
8.12.1	Adverse Events .....	38
8.12.2	Laboratory Data .....	40
8.12.3	Vital Signs.....	42



8.12.4	Electrocardiogram Data .....	43
8.12.5	Neurological Examination .....	43
8.12.6	Physical Examination.....	43
<b>9</b>	<b>SARS-COV-2 IMPACT .....</b>	<b>43</b>
<b>10</b>	<b>DATA MONITORING COMMITTEE ANALYSIS .....</b>	<b>44</b>
<b>11</b>	<b>CHANGES TO PLANNED PROTOCOL ANALYSIS .....</b>	<b>44</b>
11.1	Base Strengths (doses) of arimoclomol .....	44
11.2	Populations.....	44
11.3	Analyses.....	45
11.4	Data Collection .....	45
<b>12</b>	<b>REFERENCES .....</b>	<b>45</b>
<b>13</b>	<b>LIST OF TABLES, FIGURES AND LISTINGS .....</b>	<b>46</b>
<b>14</b>	<b>APPENDICES .....</b>	<b>61</b>
14.1	Appendix 1 .....	61
14.2	Appendix 2.....	62
14.3	Appendix 3 .....	69
14.4	Appendix 4.....	69
14.5	Appendix 5.....	70
14.6	Appendix 6.....	71



## 1 INTRODUCTION

This document details the planned statistical analyses of protocol “ORARIALS-01” study titled “A Phase 3, Randomised, Placebo-Controlled Trial of Arimoclomol in Amyotrophic Lateral Sclerosis”. The proposed analyses are based on the contents of the amended protocol, version 7.0 (dated 08-Jun-2020).

This is a multicenter, randomized, double-blind, placebo-controlled, parallel group trial to evaluate the efficacy and safety of 744 mg arimoclomol (248 mg TID) over a 76 weeks’ treatment period consisting of a combination of in-person visits and remote visits (telephone calls). An independent DMC has been established to monitor the benefit: risk.

The current protocol presents strengths of arimoclomol by weight of citrate salt. However, strengths of arimoclomol are presented throughout this document by base weight (e.g., excluding citrate salt).

For details regarding change from protocol please see [Section 11](#).

## 2 STUDY OBJECTIVES

### 2.1 Primary objective

To determine the efficacy of chronic treatment with 744 mg/day arimoclomol (248 mg t.i.d.) compared to placebo over 76 weeks in patients with ALS as assessed with CAFS

### 2.2 Secondary objectives

To evaluate the impact of 744 mg/day arimoclomol (248 mg t.i.d.) compared to placebo on:

- Disease progression as measured by change from baseline of the ALSFRS-R
- Time to PAV/tracheostomy free (PAV)/tracheostomy free survival
- Progression of respiratory dysfunction as measured by change from baseline in SVC

### 2.3 Exploratory Objectives

- To explore the potential effect of 744 mg/day arimoclomol (248 mg t.i.d.) compared to placebo on cognitive and behavioral changes.
- To evaluate the effect of 744 mg/day arimoclomol (248 mg t.i.d.) compared to placebo on health-related quality of life

### 2.4 Safety Objective



To assess the safety and tolerability of 744 mg/day arimoclomol (248 mg t.i.d) compared to placebo.

### **3 ENDPOINTS**

The endpoints listed herein are further detailed in Section 8.

#### **3.1 Primary Endpoint**

CAFS over a treatment period of 76 weeks (or end-of trial)

#### **3.2 Secondary Endpoints**

- Change from Baseline during 76 weeks in ALSFRS-R<sup>[\*]</sup>
- Time to PAV/tracheostomy/death from Baseline as assessed by
  - hazard ratio
  - median event-free survival time
- Change from Baseline during 76 weeks in percent (%) predicted SVC (PPSVC) <sup>[\*]</sup>

<sup>[\*]</sup>: *The ALSFRS-R and the % predicted SVC will in fact be analyzed on the absolute scale (the actually observed values) but reported as both as such and as a change from baseline.*

#### **3.3 Exploratory Endpoints**

- Change from baseline during 76 weeks in cognitive and behavioural changes as evaluated by ECAS<sup>[\*]</sup>
- Change from baseline in SEIQoL-DW index for the patient<sup>[\*]</sup> over a treatment period of 64 weeks
- Change from baseline in SEIQoL-DW index for the caregiver<sup>[\*]</sup> over a treatment period of 64 weeks
- Change from baseline in EQ-5D-5L VAS<sup>[\*]</sup> score over a treatment period of 76 weeks
- Number (Percent) of patient's responses to EQ-5D-5L dimensions<sup>[\*]</sup> over a treatment period of 76 weeks
- Change from Baseline to Week 76 in ALSFRS-R domains<sup>[\*]</sup>

<sup>[\*]</sup>: *Will in fact be analyzed on the absolute scale (the actually observed values) but reported as both as such and as a change from baseline.*

#### **3.4 Safety Endpoints**

- Incidence and severity of TEAEs over a treatment period of 76 weeks



- Mean and change from Baseline over a treatment period of 76 weeks in clinical safety laboratory tests, and vital signs
- Incidence of potentially clinically significant abnormalities in clinical safety laboratory tests and vital signs over a treatment period of 76 weeks
- C-SSRS over a treatment period of 76 weeks

## **4 ANALYSIS POPULATIONS, OBSERVATION PERIODS AND CENSORING**

### **4.1 Analysis populations**

The following analysis populations are planned for this trial:

- Enrolled Population: The Enrolled Population includes all patients who gave informed consent and were allocated a patient number. This population will be used for the summary of patient disposition only.
- Intent-To-Treat Population (ITT): The ITT population includes all randomized patients except for screen failure patients who were randomized in error and did not receive any investigational medicinal product (IMP). Patients in the ITT population will contribute to the evaluation ‘as randomized’.
- Modified ITT (mITT): The mITT population includes patients in the ITT Population who are not on edaravone at baseline. The mITT population will be the primary population for all efficacy analyses. Patients in the mITT population will contribute to the evaluation ‘as randomized’.
- Safety Population: The safety population includes all patients who receive any amount of IMP but excludes patients on edaravone at baseline. Patients in the Safety population will be analyzed according to treatment actually received for safety analyses. This will be the population used for all safety analyses unless otherwise specified.
- Edaravone Population: Patients in ITT who were on stable treatment (i.e. minimum of 6 months) with edaravone and receive at least one dose of IMP. Subset of displays will be repeated for this population no formal analyses, however, will be conducted.

The “actual treatment” will be based on a simple count of capsules administered from either one of the two trial treatments (e.g., treatment with largest number of capsules will be considered the actual treatment).

### **4.2 Observation Periods**

Patients and the data to be used in an analysis will be selected in a two-step manner.



Firstly, patients will be selected based on the specified analysis set.

Secondly, data points from the selected patients from the first step will be selected based on the specified observation period.

Information collected with onset date outside the observation period will be treated as missing and therefore excluded from the corresponding analysis.

Two observation periods are defined

#### **4.2.1 In-Trial**

This observation period represents the time period where patients are considered to be randomized, enrolled in the trial and as such under systematic follow-up. While the in-trial period unambiguously starts at randomization the end of the in-trial period is more involved: some patients die before the landmark visit and some withdraw informed consent (IC) prior to the landmark visit – some of which allow follow-up on the survival endpoint at the planned landmark visit and others don't allow this. Moreover, some patients completing the landmark visit “roll-over” to the open label extension trial with the week-76 visit being the baseline in the OL extension while other patients complete the current trial with or without a safety follow-up visit scheduled some 4 weeks after the landmark. Lastly, patients experiencing one of the three events comprising the composite survival endpoint are per protocol to be considered as completers and should per protocol be removed from follow-up. While certainly death and most cases of tracheostomies have indeed been treated as stipulated in the protocol, some cases of PAV have been reported with some time lag and, consequently, have not been treated as “completers” at the reported time of event.

Briefly, the in-trial period will be defined to end no later than at the week 76 visit irrespective of observations post that visit. The reason for censoring the in-trial at the week 76 visit is that follow-up post this time may be quite biased due to the selection taking place into the open label extension trial.

Due to the SARS-CoV-2 pandemic in 2020 and 2021 some study procedures have been difficult to implement as thoroughly as one might otherwise have implemented in different circumstances. Consequently, the database prior to lock and unblinding may seem ambiguous on the data points used to define the *end* of the below mentioned in-trial period. Prior to lock, a meeting will be held where the recorded data for all ITT patients (including site comments and note to file information) will be reviewed and a patient-by-patient end-date of the in-trial period set (based on the rules detailed below). The minutes of this meeting will be signed, locked and be part of the TMF. For further details see section 9.



Consequently, we define the in-trial period to start at randomization and to end at the date of below:

1. Patients recorded dead:
  - 1.1. Patients with withdrawal date: end date is the maximum of withdrawal date and all visit dates where assessments took place<sup>[1,§]</sup>
  - 1.2. Patients with completion date: end date is the minimum date of date of death, date of completion and maximum date of all visit dates where assessments took place
2. Patients withdrawing without being registered as dead or with non-lethal event: end date is the maximum of withdrawal date and all visit dates where assessments took place<sup>[1,2,§]</sup>
3. Patients completing without being registered as dead or with non-lethal event: end date is date of visit 21
4. Patients completing without being registered as dead but with non-lethal event:
  - 4.1. for patients with scheduled visits 21:
    - 4.1.1. if the last scheduled visit (either 21 or 22) is less than 76 weeks after randomization, the end date is the date of the last scheduled visit
    - 4.1.2. if the last scheduled visit (either 21 or 22) is more than 76 weeks after randomization, the end date is the date of visit 21
  - 4.2. for patients without scheduled visits 21: end date if the date completion
5. If end-of-study date is still missing after the above algorithm, the end-of-study date will be set to the date of the last recorded visit

<sup>[1]</sup>: for patients allowing follow-up of survival endpoint after withdrawal<sup>[§]</sup>, if a recorded death (or non-lethal event) lies before the 76 weeks mark, the end-date is set to this recorded date-of-death (or non-lethal event). If a death lies after the 76 weeks mark, the end-date is set to the date of post-withdrawal contact.

<sup>[2]</sup>: for patients allowing follow-up of survival endpoint after withdrawal without being recorded as dead or with a non-lethal event, the end-date is set to the date of post-withdrawal contact.

<sup>§</sup>: The protocol stipulates that for withdrawn patients, if allowed by patients or caregiver, follow-up regarding PAV/tracheostomy/death is allowed. With regard for the definition of the in-trial period, as a time period wherein patients are “under observation”, the time from date of trial withdrawal to “post withdrawal follow-up PAV/tracheostomy/death” is not relevant for say AEs, safety lab etc. Rather, this time period is only relevant for PAV/tracheostomy/death – these events are the only ones the patients are at risk for while being observed.

Consequently, for the patients allowing “post withdrawal follow-up PAV/tracheostomy/-death” (the End-of-trial eCRF page has a “For subjects that withdrew without meeting survival endpoint, did subject and/or caregiver consent to be followed up regarding survival endpoints (PAV/tracheostomy/death) at the scheduled week 76?” field), the in-trial period pertaining to PAV/tracheostomy/death, and for these endpoints only, ends at the time of what would have been the patients’ scheduled Week 76 or time of PAV/tracheostomy/death whichever came first.

It means that for the analysis of CAFS and for these survival endpoints the in-trial period may be “extended”. Specifically, the end-date of the in-trial period is defined as date of contact if the patient didn’t have an event or onset date of event if the patient had an event.



*It is important to stress that this “exception” only has to do with the aforementioned endpoints and the detailed patients*

The above *prioritized* numbered list of *checks* to find the end of the in-trial period implies that once a date “up” in the hierarchy is identified dates “further down” in the hierarchy are ignored. It also means that deaths and SAEs, say, between landmark and the safety follow-up visit will be considered as “post-trial” and only reported in lists.

Data assessed at the randomization visit (no 2) or screening data considered as baseline belongs to the in-trial period.

#### **4.2.2 On-Treatment**

This observation period represents the time period where patients are considered treated with IMP. The on-treatment period is a subset of the in-trial period and starts at the date of first administration of IMP. The protocol allows for patients to temporarily be taken off IMP by the decision of their investigator and later reinstated on IMP. As per the protocol such *drug holidays*  $\leq$  28 days are allowed.

Because the on-treatment period serves two different purposes: as a vehicle for safety reporting and as sensitivity analysis for efficacy, end of the on-treatment period will be defined slightly different for the purpose of the safety and efficacy presentations:

##### **4.2.2.1 Safety**

In order to report safety assessments during IMP interruptions in a consistent manner relative to assessments after final termination of IMP, a patient is considered to be on-treatment at any time point up to 14 days since the latest preceding administration of IMP or the end-date for the in-trial period, whichever comes first. This means that a safety assessment occurring 15 days (or later) within a (temporary) IMP interruption period will not be considered to occur on-treatment.

The definition implies that the on-treatment time for a given patient potentially may be composed by several periods separated by interruption intervals.

##### **4.2.2.2 Efficacy**

The on-treatment for the evaluation of efficacy endpoints is intended to reflect an effect that can be attributed as close as possible to a continuous treatment regimen (i.e. without interruptions that potentially can weaken the effect following an interruption).

For that reason, the on-treatment period used for efficacy evaluation ends at the first time point where a patient has not administered IMP for consecutive 15 days or more, or at or the end-date for the in-trial period, whichever comes first.



### 4.3 Observation periods and population sets

Unless otherwise specified, the primary estimand (Section 5) is associated with analyses on the mITT population and in-trial period while the secondary estimand is associated with analyses on the Safety populations and the on-treatment period.

However, for efficacy analyses planned/randomized treatments will be used whereas for safety analyses actual treatments will be used.

## 5 ESTIMANDS

In the following, estimands for the primary and secondary objectives are specified. These specifications include high-level considerations regarding either de facto missing data or data treated as missing for a specific analysis.

Two estimands are defined for the primary composite endpoint, the CAFS. An attractive feature of the CAFS is that it is meaningfully defined also for patients who die during the course of the trial. The same feature is present for the 3-composite survival endpoint.

The same feature is, however, not present for any of the other endpoints: for example, it is nonsensical to consider the ALSFRS-R score for a dead person and, likewise, a dead person has no heart rate. Because it is expected that a substantial number of the randomized patients will die during the course of the trial, the interpretation of, say, a mean change from baseline in a safety lab parameter, either becomes meaningless for the dead patients, or, rather, conditional on patients being alive at planned follow-up at 76 weeks after baseline. Therefore, two parallel “while on treatment/alive”-estimands will be defined mirroring the estimands pertaining to the primary endpoint and the survival endpoints. Further discussion below.

### 5.1 Primary Estimand

The primary estimand – “Treatment policy” is defined as follows:

- *Rank difference in the CAFS instrument between arimoclomol and control at 76 weeks for all randomized patients regardless of exposure, adherence to randomized treatment and changes in standard of care.*

The treatment policy estimand assesses the expected (rank) benefit on the CAFS instrument in a future population that results from patients being offered treatment with arimoclomol as add-on to standard-of-care as compared to standard-of-care alone.

Generalization of this estimand depends among other things on the extent to which the standard-of-care provided in this trial reflects clinical practice and whether the adherence to



trial product administration in this trial reflects the behavior of the target population. Accordingly, data collected regardless of discontinuation of trial product or background therapy will be used to draw inference.

The analysis of the primary estimand will be based on the mITT and the in-trial observation period (see caveats regarding “in-trial” definition for the CAFS Section 4.2.1).

## 5.2 Secondary Estimand

The secondary and supportive estimand – “Hypothetical” estimand, is defined as follows:

- *Rank difference of the CAFS instrument between arimoclomol and control at 76 weeks for all randomized and exposed patients if all patients adhered to treatment up until experiencing the PAV/Tracheostomy/Death endpoint or the week-76 visit.*

The hypothetical estimand assesses the rank benefit on the CAFS instrument that a future population would be expected to experience if the patients did not discontinue arimoclomol when compared to standard-of-care. It is considered a clinically relevant estimand as it provides information to treating clinicians about the expected efficacy of arimoclomol compared to standard-of-care for purposes of treating individual patients. Generalization of this estimand depends among other things on the extent to which the adherence to trial product administration in this trial could reflect the behavior of the target population. Accordingly, only data collected while patients were exposed to trial IMP will be used to draw inference.

The analysis of the primary estimand will be based on the Safety and the on-treatment observation period.

## 5.3 While on treatment/alive estimands

The two estimands defined above could in principle easily be converted to cover, say ALSFRS-R or PPSVC by replacing the CAFS with, say; PPSVC. The problem with this approach is that death is an integral part of the disease progression in ALS and as such it may be nonsensical to address a change to week 76 as many patients will not survive for that long. Rather, to accommodate this feature of the disease “while on treatment/alive”-estimands are detailed in the ICH E9 guideline on estimands. One crucial feature of this types of estimands is that the interpretation of a treatment effect is not linked to any particular time (like for example week 76).

We will define the estimands generically to apply to all the relevant efficacy parameters detailed herein. The estimands do, therefore, not address any particular endpoint like PPSVC nor specify the “unit” of the treatment comparison. The two main features of each of the estimands below are 1) the treatment contrast/comparison is not linked to any particular time



point and 2) the specifically mentioned intercurrent events used for (de)selecting data used in a particular analysis.

“De facto and while alive<sup>[x]</sup>”-estimand is defined as follows:

- *Treatment contrast in a particular parameter<sup>[\*]</sup> between arimoclomol and control during 76 weeks for all randomized patients regardless of exposure, adherence to randomized treatment and changes in standard of care.*
  - *All data for all randomized patients obtained before onset of 3-composite event, loss of follow-up (withdrawal) or before the Landmark week 76 visit will be applicable for this estimand. In other words, the first occurrence of any of these intercurrent events will terminate the follow-up on a patient for this estimand: PAV/tracheostomy/death/withdrawal/loss of follow-up/Week 76 visit*

“De jure and while alive<sup>[x]</sup> and on treatment”-estimand

- *Treatment contrast in a particular parameter<sup>[\*]</sup> between arimoclomol and control during 76 weeks for all randomized and exposed patients as long as the patients remained exposed to treatment.*
  - *All data for all randomized and treated patients obtained before onset of 3-composite event, loss of follow-up (withdrawal), date of last dose plus 14 days or before the Landmark week 76 visit will be applicable for this estimand. In other words, the first occurrence of any of these intercurrent events will terminate the follow-up on a patient for this estimand: PAV/tracheostomy/death/withdrawal/loss of follow-up/Week 76 visit/Date of last dose plus 14 days.*

[\*]: the term “particular parameter” is used to imply that a feature of the statistical model used to analyze the endpoints and the estimands is an estimate of a treatment effect. This treatment effect, can be a difference/ratio in means, a difference/ratio between regression coefficients etc.

[x]: Any event of the 3-composite survival endpoint used for the primary analysis of the CAFS (PAV/tracheostomy/death) will serve as intercurrent event.

The “while on . . .”-estimands will be analyzed using random slope- and intercepts models (see details in [Section 8.2](#)) using the entire sequence of observations pertaining to any observation period regardless of obtained as scheduled or unscheduled visits.

The analysis of the *de facto and while alive* estimand will be based on the mITT and the in-trial observation period using randomized treatment.

The analysis of the *de jure and while alive and on treatment* estimand will be based on the Safety and the on-treatment observation period using randomized treatment.



## 6 VISIT WINDOWING AND BASELINE

### 6.1 Baseline

Baseline is defined as the last non-missing value (either scheduled, unscheduled or repeat) before the patient receives the first dose of study drug (IMP).

### 6.2 Windowing

The summary and analysis of the study endpoints by visit is conditioned on each patient contributing no more than one observation per visit. As some parameters may have assessments occurring very close in time and thus may be attributable to the same planned visit, an algorithm whereby maximally 1 (one) is being selected for summary analysis is detailed below.

There are two observation periods (in-trial and on-treatment). When doing in-trial windowing, all visits are included in the process whereas for on-treatment windowing only visits at or before date of last dose+14 are included (for the latter identify the assessments that meet these criteria before windowing). It should be noted that the last assessment (e.g. end of study (EOS)) should be flagged as eligible for analysis. It may be that if last assessment falls into a window with another assessment, that the last assessment would be assigned to the next visit window (see below for rules)

For purposes of by-visit analysis, the following endpoints assessed (Vital sign, Clinical safety laboratory test, ECG, SEIQol-DW (patient; caregiver), EQ-5D-5L, SVC, C-SSRS, ALSFRS-R assessments), will be assigned to analysis visits based on the below time windows as appropriate.

For efficacy endpoints, only scheduled visits will be assigned to the defined windows (below). The analysis of all visits (including unscheduled visits) may be performed as sensitivity analyses. For efficacy endpoints, both in-trial and on-treatment windowing will be done.

For all safety endpoints, all scheduled and unscheduled visits with observations will be assigned to the defined windows (below). Only on-treatment windowing will be done for safety endpoints. If more than one measurement falls into a visit window, the measurement collected closest to the scheduled date will be used in the analysis. In case where measurements were collected an equal number of days before and after the scheduled day of the visit window then the later of the two measurements will be used for analysis. The exception is the EOS assessment, if this assessment falls into a window with another assessment and based on rules



wouldn't be flagged as eligible then EOS takes precedence and will be assigned to the next window visit and mark for analysis.

Assessments not marked "as eligible" for analysis will not be used in any analysis (descriptive or inferential) using visits as a means for organizing data and results.

All collected data will be listed and eligibility will be flagged.

### Visit Windows

			Time Window (Days)										
Scheduled			Assessments collected at in-clinic visit (except at weeks 8, 16, 24)	Assessments collected at Week 20, 52 and 76 post-baseline (only)	ALSFRS-R and CSSRS	Clinical Safety Labs	Weight						
Visit	Month	Day											
1*	Screening	-28											
2*	Baseline	1											
3*	Week 4	28	2	56			2	42	2	42			
4 <sup>r,*v5</sup>	Week 8	56					43	70	43	70			
5*	Week 12	84	57	112			71	98	71	98			
6 <sup>r,*v5</sup>	Week 16	112					99	126	99	126			
7*	Week 20	140	113	168	2	252	127	154	127	154			
8 <sup>r,*v5</sup>	Week 24	168					155	182	155	182			
9*	Week 28	196	169	224			183	210	183	224			
10 <sup>r</sup>	Week 32	224					211	238					
11*	Week 36	252	225	280			239	266	225	280			
12 <sup>r</sup>	Week 40	280					267	294					
13*	Week 44	308	281	336			295	322	281	336			
14 <sup>r</sup>	Week 48	336					323	350					
15*	Week 52	364	337	406	253	448	351	378	337	406	2	448	
16 <sup>r</sup>	Week 56	392					379	406					
17 <sup>r</sup>	Week 60	420					407	434					
18*	Week 64	448	407	490			435	462	407	490			
19 <sup>r</sup>	Week 68	476					463	490					
20 <sup>r</sup>	Week 72	504					491	518					
21*	Week 76	532	491	EOP <sup>a</sup>	449	EOP <sup>a</sup>	519	EOP <sup>a</sup>	491	EOP <sup>a</sup>	449	EOP <sup>a</sup>	

\* In-clinic/person visit: In the event a patient is no longer able to attend the site, site staff will conduct a home visit and/or by telephone call. It should be noted for home visits, certain questionnaires are collected via phone. Only a subset of clinical assessments will be conducted; those should also be included in the above windowing scheme.

<sup>a</sup>Remote Visits



<sup>v5</sup>In-person visits per protocol amendment v5

<sup>a</sup>EOP=end date (day) of in-trial period.

“Any time, Post-Baseline” visit is derived and should be based on all post-baseline assessments (scheduled/unscheduled).

## 7 SAMPLE SIZE

Two hundred and thirty-one (231) patients are planned for randomization. This number includes 213 patients according to the main trial eligibility criteria and up to an additional 18 patients on stable (i.e. minimum 6 months') treatment with edaravone and who otherwise fulfil the eligibility criteria. The sample size calculation has been made for the 213 patients in the main trial population (not including the additional patients in the edaravone sub-population).

The primary endpoint is the CAFS over a 76 weeks' treatment period comparing 744 mg/day arimoclomol (248 mg t.i.d) vs. placebo. The trial will be powered for the primary endpoint.

In the Phase 2 trial in ALS patients with pathogenic SOD1 mutation associated with rapid progression, the effect size in CAFS between arimoclomol and placebo ranged between 0.41 (entire patient population) to 0.5 (patients with baseline ALSFRS-R  $\geq 35$ ).

Further, an examination of the Pooled Resource Open-Access ALS Clinical Trials (PRO-ACT) Database reveals that the effect of active drugs versus placebo increases from 12 to 18 months on both survival and change in ALSFRS-R.

While the precise magnitude of the increase in the comparative efficacy from 12 to 18 months is difficult to assess, the rate of increase appears to be greater than a linear increase.

Applying this observation to the effect of arimoclomol compared to placebo, it is expected that the effect size (which is estimated to be at least 0.41 at 12 months) will be no less than 0.48 assuming a minor increase from 12 to 18 months, and as high as 0.61 with a linear increase from 12 to 18 months. Assuming an effect size of 0.48, 213 patients randomized 2:1 to arimoclomol and placebo will provide 90% power to detect a statistically significant difference between arimoclomol and placebo, at a two-sided type-1 error of 0.05.

The following table provides scenarios for power and sample sizes:

Effect size at 76 weeks	Power at final analysis
0.44	85%
Assumed 0.48	90%



0.54	95%
0.61	98%

The assumed effect size of 0.48 will provide 90% power with a final analysis done with a two-sided type-1 error of 0.05.

## 8 STATISTICAL ANALYSIS AND REPORTING

### 8.1 Introduction

Data processing, tabulation of descriptive statistics, calculation of inferential statistics, and graphical representations, except for PK parameter estimation, will be performed primarily using SAS (release 9.4 or higher). If the use of other software is warranted, the final statistical methodology report will detail what software was used for what purposes.

Continuous (quantitative) variable summaries will include the number of patients (n) with non-missing values, mean, standard deviation (SD), median, minimum, maximum, quartile 1 and quartile 3.

Categorical (qualitative) variable summaries will include the number of patients (n) with non-missing values, frequency and percentage of patients who are in the particular category or each possible value. In general, the denominator for the percentage calculation will be based upon the total number of patients in the trial population for the treatment groups, unless otherwise specified. The denominator for by-visit displays will be the number of patients in the relevant trial population with non-missing data at each visit.

Percentages will be based on the number of non-missing observations or the patient population unless otherwise specified. For each variable, all categories will be shown. Zero frequencies (but not the percent) within a category will be presented.

Section 6.2 describes rules for deciding which visits/assessments are eligible for (which) analysis. Please note that all data will appear in lists and for analysis non-eligible datapoints will be highlighted in the lists.

Treatment group labels will be displayed as follows: »Arimoclomol 744 mg « and »Placebo«.

The above apply to all efficacy (including single items and sub-scales) and safety endpoints of vital signs, laboratory assessments and biochemistry.

Furthermore, all summary tables of efficacy endpoints will include / be repeated for both observation periods and for absolute and change from baseline values, respectively.



Furthermore, all summary tables of safety endpoints will be performed for the on-treatment observation period and for absolute and change from baseline values, respectively.

### 8.1.1 Decimal Places

The minimum and maximum will be reported with the same degree of precision (i.e., the same number of decimal places) as the observed data.

Decimal places for derived data will be determined by the scale of measurement unless otherwise stated. No decimal places will be displayed if the smallest calculated value is  $\geq 100$ ; 1 decimal place will be displayed when the smallest value is within the interval (10, 100), with 10 being inclusive; 2 decimal places will be displayed when the smallest value is within (1, 10), with 1 being inclusive; and so on for even smaller scales of measurement.

Derived data where it is known in advance the result will be an integer for example day, month, year, number of days and total scores (for rating scales) will be presented with zero decimal places.

Means, medians and percentiles will be displayed to one more decimal place than the data, dispersion statistics (e.g. standard deviation) will have two more decimal places, and the minimum and maximum will be displayed to the same number of decimal places as reported in the raw data. Percentages will be displayed with one decimal place.

P-values will be quoted to 3 decimal places. P-values  $< 0.001$  will be presented as  $p < 0.001$ .

## 8.2 Statistical models

Unless otherwise indicated, all statistical tests will be conducted at the 0.05 significance level using 2-tailed tests, and P values will be reported. Corresponding 95% confidence intervals (CIs) will be presented for statistical tests.

Unless otherwise indicated the following model-specifications will apply to the random intercept and slope: this is a regression model of the absolute value of (assumed) continuous endpoints:

$$y_{i(j),j,t} = \underbrace{(\mu + b_{i(j),j})}_{\text{intercept}} + \underbrace{(\alpha_j + a_{i(j),j}) \cdot t}_{\text{slope}} + C_{i(j),j} + \epsilon_{i(j),j,t}$$

where the two random terms ( $b$  and  $a$ ) vary according to a 2-dimensional normal distribution with zero-mean and unstructured covariance matrix. The model includes all baseline and post baseline absolute assessments of the endpoint as dependent variables. The fixed slopes vary with treatment and the residual terms are assumed independently normally distributed with zero mean. The term  $C$  is also an intercept-term which *may* be added to the model controlling



for baseline effects (including treatment). In the model time ( $t$ ) is (actual) time in years since baseline. To ease interpretation the scale (unit) of the time covariate will be years.

The different between regression coefficients of the two treatments will be output as well as estimates of absolute and change from baseline values after 52 ( $t = 1$  in the model above) and 76 weeks ( $t = 1\frac{1}{2}$  in the model above), respectively.

To assess the adequacy of the linear model a quadratic (fixed) term will be added  $(\gamma + \gamma_j) \cdot t^2$  where  $\gamma$  represents an overall effect of squared time and the  $\gamma_j$  represents treatment specific effects of squared time. If either of these effects are significant at a 10% significance level, outputting the difference between linear slopes provides little information pertaining to treatment effects. Instead, the estimated difference between active and placebo will be outputted for time 52 weeks and 76 weeks, respectively.

### 8.3 Primary endpoint and analyses

The Combined Assessment of Function and Survival (CAFS) is the primary endpoint. It is a composite endpoint based on change from baseline in ALSFRS-R and survival endpoint (PAV, tracheostomy or death). A patient's CAFS score represents a patient's rank in the study based on comparing said patient's outcome to all other patients in the study in a pairwise fashion.

The following description is detailed. First, we detail the composite parts (the ALSFRS-R and the survival endpoints), then we outline how to "score" or "rank" the CAFS.

In Appendix 2 we provide a further discussion on details on the implementation and analyses of the CAFS.

#### 8.3.1 ALS Functioning Rating Scale – Revised (ALSFRS-R)

The ALSFRS-R is a short ordinal rating scale used to determine patients' patientive assessment of their capability and independence with 12 functional activities ('speech', 'salivation', 'swallowing', 'handwriting', 'cutting food and handling utensils', 'dressing and hygiene', 'turning in bed and adjusting bed clothes', 'walking', 'climbing stairs', 'dyspnea', 'orthopnea' and 'respiratory insufficiency').

Each activity/item is rated on a 5-point scale (from 0 to 4), giving a maximum ALSFRS-R score of 48 (sum of all 12 items). The higher the score the better functioning. Additionally, the items can be grouped into 4 functional domains consisting of 3 questions each (domain score ranges 0-12) as follows:

Domain	ALSFRS-R item number
Bulbar	Items: 1,2,3
Fine Motor	Items: 4, 5a or 5b, 6



Gross Motor	Items: 7, 8, 9
Respiratory	Items: 10, 11, 12

### 8.3.2 Time to PAV/Tracheostomy/Death

For patients who experience an event, their time-to-event will be determined from date of randomization to the date of the event (date of event-randomization date +1). Type of event met, and date of the event is captured on the end of trial CRF page. This also includes patients who withdrew prematurely from the study without having an event and consented to be contacted at what would have been their scheduled week 76 visit to find out if they had an event since withdraw. If multiple events occur (e.g., tracheostomy then death), time to first event will be used (e.g., tracheostomy). Date of PAV event is the first day of 7 consecutive days on which PAV was used for > 22 hours/day.

Patients who do not experience an event will be censored. Their censoring time will be the end of the observation period being used for analysis: for patients not allowing post-withdrawal follow-up this time date of withdrawal while for patients consenting to follow-up the time of contact. Patients who are censored will have a censoring variable, CENSORED set to 1 whereas patients who had an event will have CENSORED=0.

### 8.3.3 CAFS: Combined Assessment of Function and Survival

All analyzed patients contribute with their time of event (if relevant) and sequence of observed (observation period specific) change from baseline in ALSFRS-R. The coding/calculation of CAFS depends on the pair-wise comparison of all patients using the following algorithm:

#### 8.3.3.1 *Generic scoring algorithm*

Below we use the term “at risk” to imply that patients are “at risk *and under observation*” for the *first* of any of the event-types comprising the 3-composite survival endpoint.

- Patient S is compared to patient  $_S$ . In each patient-to-patient comparison only the index patient “S” gets scored – the comparator patient “ $_S$ ” gets scored when she (later) becomes the index patient of the algorithm
  - if S has an event at time  $t_S$  and  $_S$  is at risk at time  $t_S$ , S get a score of -1
  - if S has is at risk at time  $t_S$  and  $_S$  has an event at time  $t_{_S}$ , S get a score of +1
  - if S and  $_S$  experience events at the same time, comparison will follow the methods when comparing patients without events, noting that ALSFRS-R after event date(s) are (treated as) missing regardless of event-type
  - if none of S or  $_S$  have events, or they have coinciding events:



- find the minimum/first *time/visit number* of S's and  $\_S$ 's maximal/last value of change from baseline in ALSFRS-R<sup>[1]</sup>
- compare the thus identified "last common" change from baseline in ALSFRS-R (CfB)<sup>[2]</sup>:
- if S's CfB is smaller than  $\_S$ 's CfB then S get a score of -1
- if S's CfB is larger than  $\_S$ 's CfB then S get a score of +1
- if S's CfB is equal to that of  $\_S$ 's CfB then S get a score of 0
- Once all pairwise comparisons are made, sum the scores within each patient and denote the sum scores  $\mathcal{S}_i$ .
- The  $\mathcal{S}_i$ -scores are ranked from the lowest values (negative) to the highest (positive) of  $\mathcal{S}_i$
- The ranks are then analyzed using non-parametric methods suited for such ranks.

<sup>[1]</sup>:  $\min(\max(ts_1, ts_2, ts_3, \dots), \max(t_{\_S1}, t_{\_S2}, t_{\_S3}, \dots))$  where  $t_{ix}$  is the time (months) of i's visit x. If the two compared patients have coinciding events ALSFRS-R observations *after* the event(s) will be treated as missing.

<sup>[2]</sup>: For the calculation of CAFS based on "actual time" (see Appendix 2) there is no problem with intermittent missing data but there may indeed be problems with "how" to perform the matching of ALSFRS-R values from the two compared patients. On the other hand, when performing the matching based on "visits" the matching is trivial but here we may run into problems with intermittent missing data. Appendix 2 details different ways of handling these issues and thus some of the different sensitivity analyses planned.

### 8.3.3.2 Primary endpoint

The primary derived CAFS will use the scoring algorithm above based on all in-trial data regardless of whether obtained from scheduled or unscheduled visits and using the actual time of the assessment as "time". The included assessments are based on in-trial data (no censoring at time of last dose) and using (linear) interpolation to get matching ALSFRS-R pairs.

### 8.3.3.3 Analysis

The rank scores will be analyzed using Gehan's extended Wilcoxon's test (Gehan 1965). The test is implemented in SAS's *lifetest* procedure.

## 8.4 Sensitivity analyses of the primary endpoint

To investigate the sensitivity of the primary analysis results, complementary and separate analyses will be performed for the primary and secondary estimand. In line with European Medicines Agency (4) recommendations and with a report from the US National Research



Council (5), these analyses will primarily evaluate the sensitivity of the results due to the impact of missing data.

Moreover, [REDACTED] it was recommended not to impute survival endpoints/time for withdrawn patients not allowing follow-up on the survival endpoints after planned 76 weeks. [REDACTED] expressed some concerns regarding the CAFS as detailed above, essentially due to the (simple) censoring of patients at time of withdrawal, [REDACTED]

Lastly, as discussed in Appendix 2 there are a number of rather arbitrary choices to be made when implementing the patient-to-patient CAFS scoring. These will have to be addressed as part of the sensitivity analyses also, although not particularly associated with “missing data”.

List of *in-trial* sensitivity analyses:

- A) *Visit windows*: CAFS "as is" based on in-trial data using the in-windowed assessments of ALSFRS-R and with linear interpolation to account for intermittent missing data.
- B) *Closest match*: CAFS "as is" based on in-trial data using the assessments' actual time and using closest-match to get matching ALSFRS-R pairs.
- C) *Largest lower*: CAFS "as is" based on in-trial data using the assessments' actual time and using the use largest/latest assessment with actual time  $\leq$  the patient with the shortest follow-up's last assessment to get matching ALSFRS-R pairs.
- D) *Copy2Reference (C2R)*: for patients with follow-up on survival endpoint after withdrawal, simulation of ALSFRS-R from withdrawal to  $\min(76w, t^d)$  at scheduled time-points beginning with/at the first planned visit after withdrawal.  
Withdrawn patients without follow-up / lost to follow-up will be censored at the time of last ALSFRS-R assessment.
- E) *C2R worst case*: as detailed above (D) but where the withdrawn patients without follow-up on the survival endpoints will be treated as if they experienced an event at the time of withdrawal (last observed ALSFRS-R assessment)
- F) *C2R best case*: as detailed above (D) but where the withdrawn patients without follow-up on the survival endpoints will be treated as if they did never experience an event during the course of the trial and consequently will have their ALSFRS-R imputed as any other withdrawals up until  $\min(76w, t^d)$ .
- G) On the primary derived CAFS scores apply a semi-parametric rank analysis of covariance, adjusting for baseline ALSFRS-R and PPSVC. This is a three-step procedure akin to that outlined in Stokes *et al* (1) section 7.6. See Appendix 3 for details.

Details on the above analyses are described in Appendix 2.



For clarity, in the above analyses, patients withdrawing from the trial because of meeting non-lethal components the three-composite survival endpoints are (obviously) treated as indeed meeting the survival endpoint.

One *on-treatment* analysis will be performed:

H) As the primary analysis but where

- a. use only ALSFRS-R data obtained while on-treatment
- b. censor withdrawn patient at end-of on-treatment or withdrawal date whichever comes first (i.e. only survival-events no later than 14 days after date of last dose to be included)

Analysis where ALSFRS-R assessments not bounded by in-trial period will be performed:

During the review of blinded data of ALSFRS-R prior to unblinding, it was discovered that quite a few patients had assessments of ALSFRS-R at Visit 22 (safety follow-up) after having had a scheduled Visit 21 (Landmark at week 76) with an assessment of the ALSFRS-R.

According to protocol (version 7, 08-Jun-2020) ALSFRS-R should not be assessed at the safety follow-up and the SAP's definition of the end-of the in-trial period, for reasons discussed in detail, in general ends at the Landmark Visit 21. Potentially, for some of the affected cases, a patient may have their end-of-in-trial period end on/after V22 date, consequently, have their V22 ALSFRS-R assessments included in the in-trial period.

Although the values of the ALSFRS-R are quite consistent between the two visits and that the periods between visits are in general small, one has to assess the impact of leaving out the Visit 22 assessments for some patients. One caveat will be that for patients that entered into the open-label extension of this study. For these patients, Visit 21 (Week 76) is considered baseline and hence any 01 assessments collected after this date will not be included in the analysis.

Thus, the selection of the patients' ALSFRS-R trajectories used for the CAFS will be augmented by the possibly censored ALSFRS-R obtained after the end-date of the in-trial period.

The analysis will thus use exactly the same methodology as the primary analysis of CAFS but with "adding a few more observations on the ALSFRS-R" – potentially one for each of the subjects in the table above.

## ***Figures***



It is not obvious how one should be graphically present CAFS, but for the primary analysis two plots will be output: The first is a simple CDF-plot of the ranks (scaled to lie between 0-1) where the CDFs are estimated *within* each treatment. Secondly, a histogram of the ranks where each treatment's rank scores (scaled to lie between 0-1) get grouped into deciles.

### ***Effect size***

The CAFS analyses will be supplemented by calculating Cliff's  $\delta$ -statistic  $\frac{\#(S_{i,A} > S_{j,B}) - \#(S_{i,A} < S_{j,B})}{n_A n_C}$  for all patients  $i$  in Arimoclomol and all patients  $j$  in placebo and where  $S_{x,y}$  is the CAFS score of patient  $x$  in treatment  $y$ .

### **8.4.1 Subgroup analyses**

The primary analysis of CAFS will additionally be analyzed in subgroups (separately) based on onset of ALS »Bulbar« vs »Limb«. Patients with ALS onset in both bulbar and limb will be grouped into the »Bulbar« group.

## **8.5 Secondary endpoints and analyses**

### **8.5.1 Time to event**

Time to PAV/tracheostomy/death will be analyzed with a Cox proportional hazards model controlling for randomized treatment and (centered) baseline ALSFRS-R and PPSVC scores, respectively. Hazard ratio and 95% CI will be presented along with p-value.

Analyses will use data from both the in-trial and on-treatment observation periods, respectively, and the in-trial analyses will be carried out both including and excluding any post-withdrawal event information.

The primary analysis of interest though is the in-trial analysis including post-withdrawal event information.

In addition to the semi-parametric analyses above, the in-trial analysis including post-withdrawal event information, will be repeated using a parametric survival model assuming Weibull distributed event-times similarly controlling for randomized treatment and (centered) baseline ALSFRS-R and PPSVC scores. In this Weibull analysis, the test performed will be the estimated median event-free survival in the Arimoclomol group compared to the estimated median event-free survival in the Placebo group.

### ***Figures***

The in-trial time to first event based on all recorded events will be presented graphically using Kaplan-Meier estimates.



Two figures will be output for the Weibull analysis: 1) the integrated hazard for each treatment from the Cox-analysis and the integrated hazard for each treatment Weibull analysis are to be plotted against time since randomization. And 2) the log integrated hazard for each treatment from the Cox-analysis and the log integrated hazard for each treatment Weibull analysis are to be plotted against log time since randomization.

SAS detailed in Appendix 3.

### **8.5.2 ALSFRS-R**

For this endpoint, one analysis for each of the “*while on . .*” estimands is performed:

- Absolute ALSFRS-R will be analyzed to estimate rate of change using the random intercept and slope model previously specified. In addition to the previous specification, the regression model will include treatment as a factor and baseline centered<sup>[§]</sup> PPSVC as a covariate. The data included in the analysis include all observed observation period specific data regardless of windowing and whether the data originate from scheduled or unscheduled visits.

<sup>[§]</sup>: *individual baseline values subtracted the overall population mean at baseline.*

#### ***Figures***

Three figures of ALSFRS-R will be presented:

- A simple (in-trial) mean-plot ( $\pm$ SD) by treatment and visit corresponding to the simple summary table. Sample size per treatment and visit to be displayed.
- A (panel) plot of all individual (in-trial) trajectories using all recorded values (regardless of being in-windowed) grouped by time of last ALSFRS-R assessment. x-axis to be actual time in weeks. Use either nine or six panels and color code each treatment.
- A (panel) by-treatment plot with each subject’s estimated “regression line” is plotted overlaid the treatment (specific) regression.

#### ***Sensitivity Analysis***

Consistent with section 8.4, a sensitivity analysis using the same model as for the primary analysis of ALSFRS-R will be performed where in general all records of ALSFRS-R (from baseline and onwards) will be included irrespective of the end-date of the in-trial period.

### **8.5.3 % predicted SVC**

SVC measures the volume that can be exhaled from a full inhalation after exhaling to a maximum as slowly as possible. SVC is collected at in-clinic and at home visits. Several



attempts at the SVC are conducted and the highest SVC result from the acceptable attempts is considered the best attempt/result and used for analysis. The spirometer used in-clinic for assessments is different from the one that will be used for an at home visit. Additionally, in-clinic visit spirometry results were subject to centralized over-read analysis whereas the home visit results were not. PPSVC will be derived as follows for SVC assessments:

$$\text{PPSVC} = (\text{SVC} \div \text{predicted SVC (PSVC)}) * 100$$

where PSVC is derived per European Community of Coal and Steel (ECCS) reference equations:

- If male:  $\text{PSVC} = 0.061 \times \text{height} - 0.028 \times \text{age} - 4.65$
- If female:  $0.0466 \times \text{height} - 0.024 \times \text{age} - 3.28$ , where [height units (cm); age units (year)]

The quality of the spirometry results could impact interpretation of the results. Hence, data of “unacceptable” quality will be excluded from analysis. For in-clinic assessments, if a spirometry assessment is assigned a best test review (BTR) grade of unacceptable, then this means the data does not meet ATS criteria for data quality and hence will be excluded from analysis. Although, spirometry results from home visits are not overread and an BTR grade isn’t assigned, the spirometric device will provide “error codes” for the attempts that are used to determine the best measurement. These “error codes” represent specific issues with the measurement (e.g. error code Q means No stable baseline) which may lead to excluding the best attempt from analysis. This will depend on the error code(s) provided by the device (e.g., any attempt with error code of Q, an assessment with combination of O (no plateau) or (expiration of less than 6 sec) with error code N (no repeatability)). For PPSVC, as for ALSFRS-R, one analysis for each of the two “*while on...*” estimands is performed:

- Absolute PPSVC will be analyzed to estimate rate of change using the random intercept and slope model previously specified. In addition to the previous specification the regression model will include treatment as a factor and baseline centered ALSFRS-R as a covariate. The data included in the analysis include all observed observation period specific data regardless of windowing and whether the data originate from scheduled or unscheduled visits and regardless of whether obtained in-clinic or home-based.

Observed values and (change from baseline values) of PPSVC will only be analyzed descriptively.

### ***Figures***



Three figures akin to the ones detailed for ALSFRS-R will be presented.

### 8.5.4 Subgroup analyses

The in-trial analysis of ALSFRS-R and PPSVC will additionally be analyzed in subgroups (separately) based on onset of ALS »Bulbar« vs »Limb«. Patients with ALS onset in both bulbar and limb will be grouped into the »Bulbar« group.

## 8.6 Other efficacy endpoints

The three groups of endpoints detailed in this section will be summarized descriptively as detailed previously. Only the ECAS will be subjected to formal analysis and this just for the in-trial observation period. This analysis will be carried out using a by-visit ANCOVA which include baseline ECAS, region [pooled countries] and treatment as main effects. The least square mean estimates of change from baseline and least square mean estimates of absolute values will be reported adjusted according to observed baseline distribution. As ANCOVAs are adjusted for baseline and region, this adjustment means that estimates of change from baseline and absolute values represent predictions of a patient with baseline corresponding to the overall value of the baseline parameter and who is from an “average” region.

### 8.6.1.1 *Edinburgh Cognitive and Behavioral ALS Screen (ECAS)*

ECAS determines cognitive and behavioral changes of patients suffering from ALS. It is a global measure for the evaluation of cognitive impairments in ALS patients.

ECAS consists of 15 individual tasks, which rely on standard procedures and tests used in neuropsychology. These 15 individual tasks correlate with certain cognitive functions which are split up into 5 different subdomains of which; 3 are ALS-specific ('language', 'verbal fluency' and 'further executive functions') and 2 are non-specific functions ('memory' and 'visuospatial'). The ECAS total score of 136 points is the sum of scores of all 15 individual tasks. The lower the score the more cognitive impairment.

ECAS total Score:

ALS-Specific		ALS non-specific	
Language	Naming Comprehension Spelling	Memory	Immediate Recall Delayed retention Delayed recognition
	Fluency – letter S Fluency – letter T		Visu ospa trial Dot Counting



Executive functions	Reverse Digit Span Alternation Sentence Completion Social Cognition		Cube Counting Number location
---------------------	--	--	----------------------------------

The primary caregiver will complete the behavior screen (score 0-10) and the psychosis screen (0-3).

#### ***8.6.1.2 Schedule for the Evaluation of Individual Quality of Life-Direct Weighting (SEIQoL-DW)***

The SEIQoL-DW is a semi-structured interview for assessment of quality of life (QoL) of the individual. The interviewer first elicits the 5 areas of life (cue) considered most important by the individual in determining his/her QoL. The level of satisfaction/functioning in each area for the moment is captured by the interviewer by asking the patient to mark a bar on a scale that measures from 0 (worst possible) to 100 mm (best possible). The interviewer measures the vertical height of each bar (mm) [ each measurement is known as a cue level (mm)]. Then the interviewer asks the patient to rate how important each of the 5 areas are relative to each other in his/her life at the present [each measurement is known as a weight]. This is based on a 0 to 100 scale (%). Once this is done the weight for each cue is divided by 100 so that the final cue weights range from 0-1 (known as cue weight). An overall index called the SEIQoL-DW index (score range 0-100) is calculated as the sum of  $(\text{cue level}_i \times \text{cue weight}_i)$  where  $i$  represents the 5 cues (important areas of life).

This HRQoL interview will be administered to both the patient and their primary caregiver. Separate SEIQoL-DW Index will be calculated. If the caregiver changes in course of the study, this assessment will not be administered to the new caregiver.

#### ***8.6.1.3 EQ-5D-5L***

The EQ-5D-5L assesses a patient's health-related quality of life. The EQ-5D-5L consists of 2 components: the EQ-5D descriptive system and the health EQ visual analogue scale (EQ-VAS). The version adapted for interview over the telephone will be used in this study.

The EQ-VAS records a patient's self-rated health as a score of 0 (worst health you can imagine) to 100 (best health you can imagine).

The EQ-5D comprises the following five dimensions:

- mobility
- self-care



- usual activities
- pain/discomfort
- anxiety/depression.

Each dimension has 5 response levels ranging from no problem to extreme problem.

## 8.7 Multiplicity

To prevent type-1 error inflation, the secondary endpoints will be tested sequentially in the following order:

- ALSFRS-R: the in-trial analysis of ALSFRS-R using the random intercept and slope model
- Time to PAV/tracheostomy/death: The Cox-model applied to the in-trial observation period data including post-withdrawal information.
- Time to PAV/tracheostomy/death: The Weibull-model's estimated difference in median event-free survival applied to the in-trial observation period data including post-withdrawal information.
- PPSVC: the in-trial analysis of PPSVC using the random intercept and slope model using all data *regardless* of obtained in-clinic or at-home

This will be done at the final analysis (following statistical significance of the primary endpoint) at a two-sided 0.05 level of significance. These confirmatory tests will be based on the estimands specified in Section 5 with associated analyses in Section 8.

## 8.8 Columbia-Suicide Severity Rating Scale

The following outcomes are C-SSRS categories and have binary responses (yes/no). The categories have been re-ordered from the actual scale to facilitate the definition of composite endpoints:

Category 1	Wish to be Dead
Category 2	Non-specific Active Suicidal Thoughts
Category 3	Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act
Category 4	Active Suicidal Ideation with Some Intent to Act, without Specific Plan
Category 5	Active Suicidal Ideation with Specific Plan and Intent
Category 6	Preparatory Acts or Behavior
Category 7	Aborted Attempt
Category 8	Interrupted Attempt



Category 9	Actual Attempt (non-fatal)
Category 10	Completed Suicide

In this study, two versions of the C-SSRS is used at Screening (Baseline): lifetime assessment and recent (past 18 months). For post-baseline assessments, “Since Last Visit” version is used.

Suicidal Ideation since baseline – A “yes” answer at any time during double blind treatment to any one of the 5 suicidal ideation questions (categories 1-5) on the C-SSRS.

Suicidal Behavior since baseline – A “yes” answer at any time during double blind treatment to any one of the 5 suicidal behavior questions (categories 6-10) on the C-SSRS.

Suicidal Ideation or Suicidal Behavior- A “yes” answer at any time during double-blind treatment to any of suicidal ideation/behavior questions (categories 1-10) on the C-SSRS.

There will be no imputation of missing data for C-SSRS.

Number (percent) of patients with Suicidal Ideation, Suicidal Behavior, and Self-Injurious Behavior without Suicidal Intent Based on the C-SSRS will be summarized by treatment group and overall at any Time Post-baseline and separately for Lifetime and 18 months (baseline).

## 8.9 Other reporting issues

### 8.9.1 Derived Data

This section describes the derivations required for statistical analysis. Unless otherwise stated, variables derived in the source data will not be re-calculated.

#### *Country*

For the purpose of statistical analyses, “country” will in some cases be taking into account as a covariate. Due to, however, small sample sizes, some countries will be “merged” into “regions”. Thus, Germany, Sweden, Italy, France, Poland, and Spain are reported “as is” and the Netherlands, Belgium and Great Britain and Switzerland get merge into “Benelux, CHE and GBR” and, lastly, Canada and the United States of America get merge into “North America”.

#### *Age*

Age at time of informed consent will be used to represent age of patient in the study.

#### *Race*



Where more than one race category has been selected for a patient, these race categories will be combined into a single category labelled “Multiple Race” in the summary tables. The listings will reflect the original selected categories.

#### ***Duration/Study Day/Time***

Study day will be calculated as the number of days from first dose of investigational medicinal product (IMP).

- date of event – date of first dose of IMP + 1, for events on or after first dose
- date of event – date of first dose of IMP, for events before first dose.

#### ***Conventions for Missing and Partial Dates***

All rules explained below for partial/missing dates will be followed unless contradicted by any other data recorded on the electronic Case Report Form (eCRF).

#### ***Missing/Partial Start/Stop Date of death, tracheostomy and PAV***

No event date of any of the three events of interest are observed to less precision than “months”: If the year and month matches the end of study date and if the end of study day is less than the 15<sup>th</sup> of the month then the event date is set to end of study date, otherwise the event dates are imputed to the 15<sup>th</sup> of each month.

#### ***Missing/Partial Start/Stop Date of Adverse Events and Concomitant Medications***

Missing and partial start and stop date will be imputed for analysis purposes as follows

*Partial or missing stop date will be imputed as follows:*

If the stop date is completely missing and the event has resolved or the patient has stopped taking the concomitant medication, the stop date will be imputed as the date of the patient’s last clinic visit in the study.

- If only the year is known, the stop date will be imputed as “31-Dec” of that year or as the date of the patient’s last clinic visit in the study if in the same year.
- If the month and year are known, the stop date will be imputed as the last day of that month unless the stop date corresponds to the same month as the patient’s last clinic visit in which case the date of patient’s last clinic visit in the study will be used instead.

*Missing start date will be imputed as follows:*

- If the stop date occurs on or after the start of study drug or the event/concomitant medication is ongoing, the start date will be imputed as the date of the first dose of study drug.



- If the stop date occurs before the start of study drug, the start date of the event/concomitant medication will be imputed as the patient's screening date or the stop date of the event/concomitant medication whichever the earlier.

*Partial start date (year present, but month and day missing)*

- If the stop date occurs on or after the start of study drug or the event/concomitant medication is ongoing, and the year is the same as the year of first dosing the start date will be imputed as "01-Jan" of the same year or the date of the first dose of study drug whichever is latest. If the year is different from the year of first dosing "01-Jan" will be used.
- If the stop date occurs before the start of study drug, the start date of the event/concomitant medication will be imputed as the "01-Jan" of the same year.

*Partial start date (month and year present, but day missing)*

- If the stop date occurs on or after the start of study drug or the event/concomitant medication is ongoing, the start date will be imputed as the first day of the same month and year unless this partial start date is in same month as the first dose of study drug in which case the date of first dose of study drug will be used.
- If the stop date occurs before the start of study drug, the start date will be imputed as the first day of the month and year of the partial stop date.

If the start time is missing it will only be imputed in the case where the start date of the concomitant medication/event corresponds to the date of the first dose of study drug. The time will be imputed as the time as the first dose of study drug+ 1 min. In all other cases the time will not be imputed.

***Missing Last Date of Study Drug Dosing***

If the date of last dose of study drug is completely missing, then for analysis purposes, last dose date will be imputed as the earliest date among the following dates:

- Date patient would run out of study drug assuming full compliance from the date of the study drug was last dispensed taking into account dose (e.g., was patient de-escalated to 372 mg/day)
- Date of early withdraw or
- Date of death

Similarly, if only the month and year of last dose was recorded, then for analysis purposes, day of last dose will be imputed as the earliest day among the above dates and last day of the month of the recorded last dose.



### ***Missing Diagnosis Dates***

If the month and year are present but the day is missing, the diagnosis date will be set to first day of the relevant month. If only the year is recorded the diagnosis date will be set as “01-Jan” for that year.

### ***Study Completer***

Per protocol, a patient is followed for 76 Weeks or until they meet the survival endpoint of PAV/Tracheostomy/Death whereby once they meet survival endpoint, the patient is withdrawn from the study. Hence, a patient that completes Visit 21 (Week 76) visit or meets the survival endpoint during the study is for reporting purposes considered a completer.

### ***Exposure to Study Drug***

Exposure to IMP will be calculated as follows from the date of last dosing minus the first day of dosing + 1. The exposure calculation will not take into account breaks in therapy.

Patient-years of exposure will be calculated as the sum of the above treatment durations by treatment group for safety patients divided by 365.25.

### ***Dose De-escalation***

Per protocol, at any time during the trial, the IMP may be temporarily halted (e.g., interrupted) for up to 4 weeks for an intolerable AE. Following re-challenge at the intended dose of 744 mg/day, de-escalation from 744 mg/day (248 mg t.i.d.) to 372 mg/day (124 mg t.i.d.) may be considered. The patient will remain on this decreased dose for the remainder of the trial. A flag will be created to identify these patients.

Sites record start and stop dates of an interruption, then the sites record if dosing was re-initiated after the temporary interruption and at what dose (744 mg/day or 372 mg/day). The site also is asked if there was a de-escalation to 372 mg/day. If yes, then the site records the stop date of the 744 mg/day dose and the start date of the 372 mg/day dose.

In addition to calculating overall exposure (see above), for patients that do de-escalate, number of days exposed to 744 mg/day will be calculated as follows stop date of the 744 mg/day dose minus the first day of dosing +1. In addition, number of days exposed to 372 mg/day will be calculated as last date of 372 mg/day dose minus start date of the 372 mg/day dose + 1.

### ***Compliance***

Overall treatment compliance will be determined as follows

$$\text{Compliance} = 100 \times (\text{capsules taken}/\text{capsules expected to be taken})$$

where,



capsules taken = number of capsules expected to be taken – number of missed capsules (regardless of reason including investigator instituted temporary halt).

In this study, patients can have their dose reduced hence number of capsules expected to be taken should take into account prescribed dose.

#### ***Protocol deviation Categories***

Worldwide PD Categories will be mapped to Sponsor defined categories for presentation purposes (see Appendix 4).

#### ***Inexact Values***

In the case where a variable is recorded as “ $> x$ ”, “ $\geq x$ ”, “ $< x$ ” or “ $\leq x$ ”, a value of  $x$  will be taken for analysis purposes.

### **8.9.2 Patient Disposition**

Patient disposition will be summarized as follows:

- The number of patients in each analysis population (enrolled, ITT, mITT, Safety, Edaravone) will be tabulated by treatment group and overall
- Reasons for screen failure will be tabulated
- Number (percent) of ITT patients within country and site will be tabulated by treatment group and overall.
- The number of mITT patients that completed (on or off treatment, met survival endpoint), withdrew early along with reasons for withdrawal, for those who withdrew early without meeting a survival event, if they consented to be contacted post-withdraw to determine if they met a survival event by Week 76 and if they did will be tabulated by treatment group and overall.
- The number of mITT patients that discontinued IMP prematurely and primary reason for withdrawal will be tabulated by treatment group and overall.

### **8.9.3 Protocol Deviations**

Number (percent) of patients with Major Deviations will be presented by treatment group and Overall. Important Deviations are defined as deviations classified as critical or major.

A listing of protocol deviations will be provided within Appendix 16.2.



#### **8.9.4 Baseline Comparability**

The comparability of treatment groups with respect to patient demographics and baseline characteristics will be assessed in a descriptive manner, and no formal statistical testing will be performed for the mITT population.

Standard continuous or categorical variable summaries will be presented by randomized treatment group for the following variables

- Demographics
- General medical history
- ALS characterization/History
- El Escorial
- Body Height
- Tobacco History

#### **8.9.5 Medical History**

Separate tabulations of previous and ongoing conditions at screening will be presented by randomized treatment group and overall for mITT population. Conditions will be presented by Medical Dictionary of Regulated Activities (MedDRA) primary system organ class and preferred term.

#### **8.9.6 Prior and Concomitant Medications**

Separate tabulations will be produced for prior and concomitant medications presented by randomized treatment group and overall for the mITT population. Prior medications are defined as all medications starting before the date of first dose of IMP. Concomitant medications are defined as medications taken on or after the date of first dose of IMP. Concomitant medications will be summarized using ATC Level 2.

#### **8.9.7 Exposure to Study Drug**

Extent of exposure (number of days of exposure to IMP and patient-years of exposure) and mode of administration will be presented by treatment group and overall for the Safety population.

The number (percent) of patients that fall into the following “number of interruptions” categories (0, 1, 2, >3 interruptions). Duration of longest interruption of IMP and last dose received will be summarized by treatment group and overall for the Safety population.



## 8.9.8 Treatment Compliance

Overall compliance will be presented by treatment group and overall for the Safety population descriptively (n, mean, standard deviation, median, minimum and maximum) and number (percent) of patients in the following compliance categories (<50%, 50 to <80%, 80 to 100%, >100%).

## 8.10 Pharmacokinetic Analyses

A PK analysis plan will be defined, and data reported separately.

## 8.11 Biomarker Analyses

A Biomarker analysis plan will be created, and data reported separately.

## 8.12 Safety Analyses

Unless otherwise stated, Safety summaries are based on the on-treatment period, however, all data will be listed in the appropriate listings.

### 8.12.1 Adverse Events

A »*pre-treatment AE*« is defined as any AE that has an onset prior to the first dose of IMP.

A »*treatment emergent adverse event*« (TEAE) is defined as

- Any AE that has an onset on or after the date of first dose of IMP or
- Any pre-existing AE that has worsened in severity after the first dose of IMP.

A treatment-related AE is defined as an AE that is being assessed as probably or possibly related to the IMP. If an AE has missing relationship it is assumed to be related to the IMP and assigned as “probably related” to the IMP for analysis purposes.

Maximum severity will be assumed for an AE with missing severity.

An on-treatment period (Section 4.2.2.1) TEAE is defined as

- Any TEAE that has an onset in the on-treatment period as defined for safety (see section [4.2.2.1](#)).

An in-trial period TEAE is defined as

- Any TEAE that has an onset during the in-trial period (see Section [4.2.1](#))

Maximum severity will be assumed for an AE with missing severity.

The following tables will be presented for AEs:



- Overall incidence and the number of Pre-treatment AEs, TEAEs (on-treatment and in-trial period), Serious TEAE (SAEs; on-treatment and in-trial period), on-treatment period treatment related TEAEs, on-treatment period treatment related SAEs, on-treatment period TEAEs leading to IMP withdrawal, on-treatment period TEAEs leading to IMP interruption, on-treatment TEAEs leading to dose reduction.
- TEAE by system organ class and preferred term, incidence and number of events
- TEAE by preferred term in decreasing incidence
- TEAEs > 10% in any treatment group by preferred term in decreasing incidence

Treatment related TEAE by system organ class and preferred term, incidence and number of events

- Serious TEAE by system organ class and preferred term, incidence and number of events
- Treatment related Serious TEAE by system organ class and preferred term, incidence and number of events
- TEAE by system organ class, preferred term and maximum severity, incidence
- TEAE by system organ class, preferred term and relationship, incidence
- TEAEs leading to IMP withdrawal by system organ class and preferred term, incidence and number of events
- TEAEs leading to IMP Interruption by system organ class and preferred term, incidence and number of events
- TEAEs leading to Dose Reduction by system organ class and preferred term, incidence and number of events

Additionally, two (TEAE) displays will be produced based the following standardized MedDRA Queries (SMQs)

- Liver-Related Investigations, Signs and Symptoms (SMQ)
- Drug Related Hepatic Disorders, Severe Events only (SMQ)

Incidence, number of events and number of events per 100 patient-years will be summarized.

- In-trial period
  - TEAE by system organ class and preferred term, incidence and number of events



- Serious TEAE by system organ class and preferred term, incidence and number of events
- Listing of (all) AEs
- Listing of Serious AEs

Overall incidence of deaths and a listing of all deaths will be presented. Deaths are AEs with reported outcome of Fatal or a reported death due to disease progression (latter would not be captured as an AE).

### **8.12.2 Laboratory Data**

Individual lab parameters will be summarized descriptively by treatment and visit. Descriptive statistics will be presented for each hematology, urinalysis and serum chemistry parameter. In addition to the summaries detailed previously, geometric mean and associated CV<sup>[\$]</sup> of the ratio-to-baseline and observed value of each lab parameter for post-baseline will be output.

<sup>[\$]</sup>: The geometric mean is calculated as the exponentiated mean of the log-transformed parameter where parameter is either observed value or ratio to baseline. The associated CV is calculated as  $100\sqrt{\exp(\sigma^2) - 1}$  where  $\sigma^2$  is the variance of the log transformed parameter of interest.

Each measurement (continuous data) will be classed as below (low), within (normal), or above (high) normal range, based on ranges supplied by the laboratory used. Number (%) of patients that fall out-side and within range will be reported by visit.

Listings of patients with parameter results considered to be clinically significantly abnormal by investigator will be presented. The listings will list all lab parameter results where at least one value is considered clinically significantly abnormal.

Furthermore, two “specific” lists will be presented:

- a list of all (in-trial) AEs in patients with ALAT or ASAT  $\geq 3 \times$  ULN (in-trial).
- a list of all (in-trial) AEs in patients fulfilling criteria {ALAT  $\geq 3 \times$  ULN *and* eosinophilia<sup>[\*]</sup>  $> 5\%$ } *or* {ASAT  $\geq 3 \times$  ULN *and* eosinophilia  $> 5\%$ } (in-trial).

<sup>[\*]</sup>: SDTM.LB\$LBTEST="Eosinophils/Leukocytes"

In addition, a summary of creatinine over time (continuous) will be presented according to concomitant use of MATE/OCT-02 medications at any time (see Appendix for list of concomitant medications).

Further, number (%) of patients that have potentially clinically significant (PCS) laboratory values overall and by visit will be presented as specified below.



The following categorical lab parameters of interest with defined PCS values ([Appendix 6](#)) will be summarized descriptively by visit and overall, respectively.

***Liver function Categories of Potentially Clinical Significance<sup>\*/</sup>***

- ALAT (Alanine Aminotransferase) or ASAT (Aspartate Aminotransferase)  $\geq 3 \times$  ULN
- ALAT or ASAT  $\geq 5 \times$  ULN
- ALAT or ASAT  $\geq 8 \times$  ULN
- ALAT or ASAT  $\geq 20 \times$  ULN
- Total Bilirubin  $\geq 2 \times$  ULN
- ALP  $\geq 2 \times$  ULN
- ALP  $\geq 3 \times$  ULN
- (ALAT or ASAT  $\geq 3 \times$  ULN) & Total Bilirubin  $\geq 2 \times$  ULN & ALP  $\leq 1.5 \times$  ULN
- ALAT  $\geq 3 \times$  ULN
- ALAT  $\geq 5 \times$  ULN
- ALAT  $\geq 8 \times$  ULN
- ALAT  $\geq 20 \times$  ULN
- ALAT  $\geq 3 \times$  ULN & Total Bilirubin  $\geq 2 \times$  ULN & ALP  $\leq 1.5 \times$  ULN
- ASAT  $\geq 3 \times$  ULN
- ASAT  $\geq 5 \times$  ULN
- ASAT  $\geq 8 \times$  ULN
- ASAT  $\geq 20 \times$  ULN
- ASAT  $\geq 3 \times$  ULN & Total Bilirubin  $\geq 2 \times$  ULN & ALP  $\leq 1.5 \times$  ULN

[\*]: The single-parameter categories, like e.g. “ALAT  $\geq 3 \times$  ULN” is simply a categorization of the individual recorded value (AVAL in CDISC terminology). For “by visits” summaries, the data will be “in-windowed” records, whereas for “overall” summaries also “not in-windowed” records are eligible for summary.

For multi-parameters categories like e.g. “ALAT or ASAT  $\geq 3.0 \times$  ULN & Total Bilirubin  $\geq 2 \times$  ULN & ALP  $\leq 1.5 \times$  ULN”, because of windowing being parameter specific, it is important that parameters get compared/merged *not* by visits (after relocation/windowing) but based on the original sample dates (in CDISC terminology either “VISITNUM” or “ADT”).

The above categorical parameters are to be presented in one table for “by visit” and one table for “overall” summaries, respectively.

Furthermore, the following categorical summaries will be summarized descriptively overall:

- ALAT or ASAT  $\geq 5 \times$  ULN for more than 14 days<sup>[\$]</sup>
- (ALAT or ASAT  $\geq 3 \times$  ULN) & (Total Bilirubin  $\geq 2 \times$  ULN)



- (ALAT or ASAT  $\geq 3 \times \text{ULN}$ ) and with ongoing eosinophilia  $> 5\%$  (in SDTM.LB\$LBTEST = "Eosinophils/Leukocytes")

[§]: either or both "ALAT  $\geq 5 \times \text{ULN}$  for 14 days" and/or "ASAT  $\geq 5 \times \text{ULN}$  for 14 days".

The above categorical parameters are to be presented in one table for "overall" summary.

The following creatinine related categorical summaries will be summarized both by visit and overall, respectively:

#### ***Creatinine Categories of Potentially Clinical Significance***

- Post baseline value  $\geq 1.5 \times \text{baseline value}$
- Post-baseline value  $\geq 2 \times \text{baseline value}$
- Post-baseline value  $\geq 3 \times \text{baseline value}$

The above categorical creatinine parameters are to be presented in one table for "by visit" and one table for "overall" summaries, respectively.

#### ***Figures***

**eDISH plot:** a scatter plot based on all (both in-windowed and not in-windowed) "on-treatment" pair (identified as from the same sampling date) of bilirubin (y-axis) vs ALAT (x-axis) where each parameters has been scaled (divided) by its respective ULN-value(s). The x-axis (ALAT) will have a vertical line placed at "x = 3" (corresponding to " $3 \times \text{ULN}$ ") and the y-axis (bilirubin) a vertical line located at "y = 2" (corresponding to " $2 \times \text{ULN}$ "). Both axes to be displayed as  $\log_{10}$ .

**Spaghetti plots:** For all patients with ALAT or ASAT  $\geq 3 \times \text{ULN}$ , plots of all individual ALAT and ASAT on-treatment trajectories, respectively. Overlaid should be included the geometric mean corresponding to the data presented in tabular summaries with a band indicating the SD calculated using the delta-method approximation: Geometric mean  $\times$  SD of the log-transformed values.

### **8.12.3 Vital Signs**

Descriptive statistics for observed values and changes from baseline in the following vital signs will be presented by treatment group and visit:

- Systolic blood pressure (mmHg)
- Diastolic blood pressure (mmHg)
- Pulse rate (bpm)
- Respiration rate (breath/min)
- Body temperature (degrees Celsius)



- Body weight (kg)

In addition, the Number (%) of patients that have potentially clinically significant vital sign values overall and by visit will be presented:

#### ***Vital Sign Parameters Potentially Clinically Significant Values***

Parameters	Unit	Reference Range	PCS LOW	PCS HIGH
Pulse rate	Beats/min	60-100	< 50 and decrease of $\geq 15$	>120 and increase of $\geq 15$
Diastolic Blood Pressure	mmHg	60-90	$\leq 50$ and decrease of $\geq 15$	> 105 and increase of $\geq 15$
Systolic Blood Pressure	mmHg	100-140	$\leq 90$ and decrease of $\geq 20$	> 180 and increase of $\geq 20$
Weight	Kg		decrease of $\geq 7\%$	increase of $\geq 7\%$

#### **8.12.4 Electrocardiogram Data**

Summary tables in relation to the overall interpretation (Normal, Abnormal NCS, and Abnormal CS) by visit will be presented.

#### **8.12.5 Neurological Examination**

All data will be listed only.

#### **8.12.6 Physical Examination**

All data will be listed only.

### **9 SARS-COV-2 IMPACT**

The SARS-CoV-2 pandemic of 2020 and 2021 affected trial conduct in several ways. The most obvious are “conversions” of planned in-clinic visits to “phone-visits”. For invasive procedures like lab sampling the impact is missing data while the assessment of ALSFRS-R would be conducted via phone. The scale is, however, validated for remote assessments and the impact of the pandemic therefore is or should be small. For SVC, things are a bit different as the apparatus used at sites differ from the apparatus used for home spirometry. While all the apparatus used have been calibrated and are assumed to measure the different lung function parameters adequately, the home spirometers do not estimate predicted values and, consequently, not percent-predicted values of SVC. The in-clinic apparatus does this. For the home-based assessments of SVC, therefore, the (percent-)predicted values are calculated using standard equations based on sex and age.



Another impact of the pandemic worth noting in this SAP are some ambiguities in the information collected typically on the end-of-trial form. The end-of-trial form has several different date-fields like “date of last contact”; “date of completion”; “date of death” etc. Together with the dates recorded for the held (both scheduled and unscheduled) visits these dates constitutes “milestones” by which one can determine the fate and in-trial eligibility of each patient. Due to stress under which the trial’s last year was conducted, a few patients’ “milestones dates” appear somewhat inconsistent. A few examples are appropriate (with the caveat that the below examples are based on data pre-lock possibly with some outstanding queries):

- Patient 25002 has recorded last contact on 06/07/2020. There’s no completion date; no withdrawal date; no 3-composite survival event date and the last visit took place at 05/05/2020. This visit is recorded on the V21 eCRF page (which indicate that the patient withdrew IC as site used the V21 eCRF page at withdrawal visits).
- Patient 24015 has completion date on the 09/07/2020 but with a last contact some three months prior to this on the 08/04/2020. Patient’s last visit is recorded at the 08/04/2020 (= last contact) which is V22 without a preceding V21. The patient has a recorded PAV at the 15/04/2020 and later a death at the 03/09/2020. While, for the statistical analysis the April PAV is the index event, based alone on the *dates* it is not clear how to interpret the sequence of trial milestones.

Because of such data discrepancies prior to DBL and unblinding, a meeting will be held where the end-date of each patients’ in-trial period will be decided upon. The algorithm detailed in section 4.2 will be applied.

## **10 DATA MONITORING COMMITTEE ANALYSIS**

Data monitoring committee (DMC) analyses are described in DMC statistical analysis plan.

## **11 CHANGES TO PLANNED PROTOCOL ANALYSIS**

### **11.1 Base Strengths (doses) of arimoclomol**

SAP presents arimoclomol base strengths instead of arimoclomol citrate strengths per Protocol.

Based on sponsor provided information, 124 mg arimoclomol is equivalent to 200 mg arimoclomol citrate. Hence, the following equivalents were determined:

- 1200 mg arimoclomol citrate (400 mg TID) = 744 mg arimoclomol (248 mg TID)
- 600 mg arimoclomol citrate (200 mg TID) = 372 mg arimoclomol (124 mg TID)

### **11.2 Populations**



ITT population definition was changed from “including all randomized patients, excluding patients on edaravone, who received at least one dose of IMP” to including all randomized patients.

A modified ITT population was defined as all patients in the ITT Population who are not on edaravone. The mITT population will be the primary population for all efficacy analyses instead of the ITT population.

Safety population was defined as all randomized patients who have had at least one dose of IMP. For safety analyses, patients will be analyzed based on treatment received instead of assigned. The protocol specifies ITT as the primary population for safety.

PP population and supportive analyses based on this population have been dropped. Other supportive analyses will be done with respect to the primary and secondary endpoints.

### **11.3 Analyses**

Randomization was stratified by riluzole use (yes or no). Majority of analyses included riluzole use as a factor in the model. However, randomization has been completed and it has been determined that majority of randomized patients (>90%) are on riluzole hence due to small number of patients not taking riluzole, riluzole use has been dropped as a factor from analysis models.

The primary analysis of CAFs was changed from using a nonparametric rank analysis of covariance to Gehan’s extended Wilcoxon test (see section [8.3.3.3](#)). The nonparametric rank of analysis will be conducted as a sensitivity analysis.

Clarified that the secondary endpoint, SVC, in the protocol was % predicted SVC (PPSVC).

The primary analyses for ALSFRS-R and % predicted SVC is a random intercept- and slope-analysis rather than the protocol specified analysis of covariance (ANCOVA) or .

### **11.4 Data Collection**

Per protocol, SEIQ-DW (subject and caregiver) questionnaire was scheduled to be administered at all in-clinic visits including end of trial (Visit 21 [Week 76/ET]), however, the schedule of trial procedures table (table 1) in the protocol did not clearly show this. Thus, unfortunately, the questionnaire was not collected at Visit 21.

## **12 REFERENCES**

SAS Institute Inc., Cary, NC, 27513, USA



Stokes, ME., Davis, DS. and Koch, GG. Categorical Data Analysis Using SAS, Third Edition. SAS Institute Inc., Cary, North Carolina, USA. 2012.

Little RJA, Rubin DB. Statistical analysis with missing data: New York: John Wiley & Sons. 1987.

O'Kelly, K and Ratitch, B. Clinical Trials with Missing Data: A Guide for Practitioners. Wiley. 2014.

Committee for Medicinal Products for Human Use (CHMP). Guideline on Missing Data in Confirmatory Clinical Trials. EMA/CPMP/EWP/1776/99 Rev. 1. 2 Jul 2010.

National Academy of Sciences (NAS). The Prevention and Treatment of Missing Data in Clinical Trials. Washington D.C.: The National Academies Press. 2010.

Gehan, E. (1965) A Generalized Wilcoxon Test for Comparing Arbitrarily Singly-Censored Samples. *Biometrika*, 52, 203-223.

Quanjer Ph.H, Tammeling G.J, Cotes J.E., Perdersen O.F., Peslin R. & Yernault J-C. (1993), Lung Volumes and Forced Ventilatory Flows. *European Respiratory Journal*, 6 (suppl 16), 5-40.

### **13 LIST OF TABLES, FIGURES AND LISTINGS**

The following table includes details of the tables, figures and listings to be included within each section of the eCTD. The eCTD section is shown in bold. The following validation methods maybe used

- Independent programming of numbers and manual review of format (IP)
- Independent programming by statistician of numbers and manual review of format (Stat IP)
- Manual review (MR)
- Code review (CR)

Table Number	Table Title	Validation Method	Shell Number (if repeat)
Items in bold are not table titles but references to the section headings within eCTD.			
14.1	<b>Demographics Data</b>		
14.1.1	<b>Disposition</b>		



Table Number	Table Title	Validation Method	Shell Number (if repeat)
14.1.1.1	Patient Disposition, Analysis Populations-Enrolled Population-In-trial period	IP	
14.1.1.2	Patient Disposition, Screen Failures-Enrolled Population-In-trial period	IP	
14.1.1.3	Patient Disposition, Early Withdrawals-mITT Population-In-trial Period	IP	
14.1.1.3.1	Patient Disposition, Early Withdrawals-Edaravone Population-In-trial Period	IP	14.1.1.3
14.1.1.4	Patient Disposition, Premature Discontinuation of IMP-mITT Population-In-trial Period	IP	
14.1.1.4.1	Patient Disposition, Premature Discontinuation of IMP-Edaravone Population-In-trial Period	IP	14.1.1.4
14.1.1.5	Patient Disposition, ITT population by Country and site-ITT Population-In-trial Period	IP	
14.1.1.6	Patient Disposition, mITT population by Country and site-mITT Population-In-trial Period	IP	
14.1.1.7	Patient Disposition, Safety population by Country and site-Safety Population-In-trial Period	IP	14.1.1.6
14.1.1.8	Patient Disposition, Major Protocol Deviations-mITT Population-In-trial Period	IP	
14.1.1.9	In Clinic Post-baseline Visits Switched to remote or home visit or not done due to COVID-19 Pandemic-mITT Population-In-trial Period	IP	
14.1.2	<b>Demographics</b>	IP	
14.1.2	Demographics-mITT Population-In-trial Period	IP	
14.1.2.1	Demographics-Edaravone Population-In-trial Period	IP	14.1.2
14.1.3	<b>Baseline Characteristics</b>	IP	
14.1.3.1	Prior Medical History-mITT Population-In-trial Period	IP	
14.1.3.2	On-Going Medical History-mITT Population-In-trial Period	IP	14.1.3.1
14.1.3.3	ALS Characterization-mITT Population-In-trial Period	IP	



Table Number	Table Title	Validation Method	Shell Number (if repeat)
14.1.3.3.1	ALS Characterization-Edaravone Population-In-trial Period	IP	14.1.3.3
14.1.3.4	Tobacco History-mITT Population-In-trial Period	IP	
<b>14.2</b>	<b>Efficacy Data</b>		
<b>14.2.1</b>	<b>Primary Efficacy Endpoint</b>		
14.2.1.1	CAFs (Rank Score)-Primary Analysis-mITT Population-In-trial Period	STAT IP	
14.2.1.2	CAFs (Rank Score)- Visit Window Sensitivity Analysis-mITT Population-In-trial Period	STAT IP	14.2.1.1
14.2.1.3	CAFs (Rank Score)- Closest Match Sensitivity Analysis-mITT Population-In-trial Period	STAT IP	14.2.1.1
14.2.1.4	CAFs (Rank Score)- Largest lower Sensitivity Analysis-mITT Population-In-trial Period	STAT IP	14.2.1.1
14.2.1.5	CAFs (Rank Score)- C2R Sensitivity Analysis-mITT Population-In-trial Period	STAT IP	
14.2.1.6	CAFs (Rank Score)- C2R Worst Case Sensitivity Analysis-mITT Population-In-trial Period	STAT IP	14.2.1.5
14.2.1.7	CAFs (Rank Score)- C2R Best Case Sensitivity Analysis-mITT Population-In-trial Period	STAT IP	14.2.1.5
14.2.1.8	CAFs (Rank Score)- nonparametric rank analysis of covariance Sensitivity Analysis-mITT Population-In-trial Period	STAT IP	
14.2.1.9	CAFs (Rank Score)- on-treatment Period Sensitivity Analysis-Safety Population-On-treatment Period	STAT IP	14.2.1.1
14.2.1.10	CAFs (Rank Score)- Limb-onset ALS-mITT Population with limb-onset ALS-In-trial Period	STAT IP	14.2.1.1
14.2.1.11	CAFs (Rank Score)-Bulbar-onset ALS-mITT Population with Bulbar-onset ALS-In-trial Period	STAT IP	14.2.1.1
14.2.1.12	CAFs (Rank Score) – ALSFRS-R Sensitivity	STAT IP	14.2.1.1
<b>14.2.2</b>	<b>Secondary Efficacy Endpoints</b>	STAT IP	
14.2.2.1	Time to PAV/Tracheostomy/Death - Primary analysis (In-trial Period including post-withdrawal events)-mITT Population-In-trial Period	STAT IP	



Table Number	Table Title	Validation Method	Shell Number (if repeat)
14.2.2.2	Time to PAV/Tracheostomy/Death - Sensitivity analysis (In-trial Period excluding post-withdrawal events)-mITT Population-In-trial Period	STAT IP	14.2.2.1
14.2.2.3	Time to PAV/Tracheostomy/Death - Sensitivity analysis (On-treatment Period)-Safety Population-On-treatment Period	STAT IP	14.2.2.1
14.2.2.4	Time to PAV/Tracheostomy/Death - Weibull Analysis (In-trial Period including post-withdrawal events)-mITT Population-In-trial Period	STAT IP	
14.2.2.5	Slope analysis of ALSFRS-R Score-De Facto and while alive-mITT Population-In-trial Period	STAT IP	
14.2.2.6	Slope analysis of ALSFRS-R Score- De jure and while alive and on-treatment-Safety Population-On-treatment Period	STAT IP	14.2.2.5
14.2.2.7	ALSFDRS-R, Descriptive Summary by Visit (De facto and while alive)-mITT Population-In-trial Period	IP	
14.2.2.8	Slope analysis of % Predicted SVC- De facto and while alive-mITT Population-In-trial period	STAT IP	14.2.2.5
14.2.2.9	Slope analysis of % Predicted SVC- De jure and while alive and on treatment-Safety Population-On-treatment period	STAT IP	14.2.2.5
14.2.2.10	% Predicted SVC, Descriptive Summary by Visit (De facto and while alive)-mITT Population-In-trial Period	IP	14.2.2.7
14.2.2.11	Slope analysis of ALSFRS-R Score-De Facto and while alive- Limb onset ALS-mITT Population with limb onset ALS-In-trial Period	STAT IP	14.2.2.5
14.2.2.12	Slope analysis of ALSFRS-R Score-De Facto and while alive- Bulbar onset ALS-mITT Population with Bulbar onset ALS-In-trial Period	STAT IP	14.2.2.5
14.2.2.13	Slope analysis of % Predicted SVC-De Facto and while alive- Limb onset ALS-mITT Population with limb onset ALS-In-trial Period	STAT IP	14.2.2.5
14.2.2.14	Slope analysis of % Predicted SVC-De Facto and while alive- Bulbar onset ALS-mITT Population with Bulbar onset ALS-In-trial Period	STAT IP	14.2.2.5
14.2.2.15	Slope Analysis of ALSFRS-R Score-Sensitivity	STAT IP	14.2.2.5
<b>14.2.3</b>	<b>Exploratory Endpoints</b>		



Table Number	Table Title	Validation Method	Shell Number (if repeat)
14.2.3.1	ECAS, Descriptive Summary by Visit-mITT Population-In-trial Period		14.2.2.7/14.3.4.1
14.2.3.2	ECAS-Caregiver Screens, Descriptive Summary by Visit-mITT Population-In-trial Period		14.2.2.7/14.3.4.1
14.2.3.3	Change from baseline to Week 76/ET in ECAS Total Score- De facto and while alive-mITT Population-In-trial Period	STAT IP	14.2.2.4
14.2.3.4	SEIQoL-DW (Subject), Descriptive Summary by Visit-mITT Population-In-trial Period		14.2.2.7
14.2.3.5	SEIQoL-DW (Caregiver), Descriptive Summary by Visit-mITT Population-In-trial Period		14.2.2.7
14.2.3.6	EQ-5D-5L, Categorical Summary of Dimensions by Visit-mITT Population-In-trial Period		
14.2.3.7	EQ-5D-5L, Descriptive Summary of VAS Score by Visit-mITT Population-In-trial Period		14.2.2.7
14.2.3.8	ALSFRS-R, Categorical Summary of Item Responses by Visit-mITT Population-In-trial Period		14.2.3.6
14.2.3.9	Slope analysis of ALSFRS-R Bulbar Domain- De facto and while alive-mITT Population-In-trial period	STAT IP	14.2.2.5
14.2.3.10	Slope analysis of ALSFRS-R Fine Motor Domain- De facto and while alive-mITT Population-In-trial period	STAT IP	14.2.2.5
14.2.3.11	Slope analysis of ALSFRS-R Gross Motor Domain- De facto and while alive-mITT Population-In-trial period	STAT IP	14.2.2.5
14.2.3.12	Slope analysis of ALSFRS-R Respiratory Domain- De facto and while alive-mITT Population-In-trial period	STAT IP	14.2.2.5
14.2.3.13	ALSFRS-R Domains, Descriptive Summary by Visit (De facto and while alive)-mITT Population-In-trial period	IP	14.2.2.7/14.3.4.1
<b>14.3</b>	<b>Safety Data</b>		
<b>14.3.1</b>	<b>Displays of Adverse Events</b>		
14.3.1.1	Adverse Events, Overall Summary of Adverse Events- Safety Population-In-trial period	IP	
14.3.1.1.1	Adverse Events, Overall Summary of Adverse Events- Edaravone Population	IP	14.3.1.1



Table Number	Table Title	Validation Method	Shell Number (if repeat)
14.3.1.1.2	Adverse Events, Overall Summary of Adverse Events-Safety Population + Edaravone Population-In-trial Period		
14.3.1.2	Adverse Events, Pre-treatment by Primary System Organ Class and Preferred Term-Safety Population-In-trial Period	IP	
14.3.1.2.1	Adverse Events, Pre-treatment by Primary System Organ Class and Preferred Term-Edaravone Population-In-trial Period		
14.3.1.3	Adverse Events, TEAEs by Primary System Organ Class and Preferred Term-Safety Population-On-treatment Period	IP	14.3.1.2
14.3.1.3.1	Adverse Events, TEAEs by Primary System Organ Class and Preferred Term-Edaravone Population-On-treatment Period	IP	14.3.1.2
14.3.1.4	Adverse Events, TEAEs by Preferred Term in decreasing frequency-Safety Population-On-treatment Period	IP	
14.3.1.4.1	Adverse Events, TEAEs >10% in any treatment group-Safety Population-On-treatment Period		
14.3.1.5	Adverse Events, Treatment Related TEAEs by Primary System Organ Class and Preferred Term-Safety Population-On-treatment Period	IP	14.3.1.2
14.3.1.6	Adverse Events, Serious TEAEs by Primary System Organ Class and Preferred Term-Safety Population-On-treatment Period	IP	14.3.1.2
14.3.1.6.1	Adverse Events, Serious TEAEs by Primary System Organ Class and Preferred Term-Edaravone Population-On-treatment Period	IP	14.3.1.2
14.3.1.7	Adverse Events, Treatment Related Serious TEAEs by Primary System Organ Class and Preferred Term-Safety Population-On-treatment Period	IP	14.3.1.2
14.3.1.8	Adverse Events, TEAEs Leading to IMP Withdrawal by Primary System Organ Class and Preferred Term-Safety Population-On-treatment Period	IP	14.3.1.2
14.3.1.9	Adverse Events, TEAEs Leading to IMP Interruption by Primary System Organ Class and Preferred Term-Safety Population-On-treatment Period	IP	14.3.1.2



Table Number	Table Title	Validation Method	Shell Number (if repeat)
14.3.1.10	Adverse Events, TEAEs Leading to Dose Reduction by Primary System Organ Class and Preferred Term-Safety Population-On-treatment Period	IP	14.3.1.2
14.3.1.11	Adverse Events, TEAEs by Primary System Organ Class, Preferred Term and Maximum Severity-Safety Population-On-treatment Period	IP	
14.3.1.12	Adverse Events, TEAEs by Primary System Organ Class, Preferred Term and Relationship-Safety Population-On-treatment Period	IP	
14.3.1.13	Adverse Events, TEAEs by Primary System Organ Class and Preferred Term-Safety Population-In-trial Period	IP	14.3.1.2
14.3.1.13.1	Adverse Events, TEAEs by Primary System Organ Class and Preferred Term-Edaravone Population-In-trial Period	IP	14.3.1.2
14.3.1.14	Adverse Events, Serious TEAEs by Primary System Organ Class and Preferred Term-Safety Population-In-trial Period	IP	14.3.1.2
14.3.1.14.1	Adverse Events, Serious TEAEs by Primary System Organ Class and Preferred Term-Edaravone Population-In-trial Period	IP	14.3.1.2
14.3.1.15	Adverse Events, TEAEs - Liver-Related Investigations, Signs and Symptoms (SMQ)-Safety Population-On-treatment Period	IP	
14.3.1.16	Adverse Events, TEAEs - Drug Related Hepatic Disorders, Severe Events only (SMQ)-Safety Population-On-treatment Period	IP	
<b>14.3.2</b>	<b>Listings Of Deaths, Other Serious And Significant Adverse Events</b>		
14.3.2.1	Deaths, Listing-Enrolled Population-In-trial Period	IP	
14.3.2.2	SAE, Listing-Enrolled Population-In-trial Period	IP	
<b>14.3.3</b>	<b>Narratives of Deaths, Other Serious and Certain Other Significant Adverse Events</b>		
<b>14.3.4</b>	<b>Abnormal Laboratory Values</b>		
14.3.4.1	Hematology, Descriptive Summary of Parameters by Visit-Safety Population-On-treatment Period	IP	



Table Number	Table Title	Validation Method	Shell Number (if repeat)
14.3.4.1.1	Hematology, Descriptive Summary of Parameters by Visit-Edaravone Population-On-treatment Period	IP	14.3.4.1
14.3.4.2	Hematology, Categorical Summary of Potentially Clinically Significant Values at Any Time, Post-Baseline-Safety Population-On-treatment Period	IP	
14.3.4.2.1	Hematology, Categorical Summary of Potentially Clinically Significant Values at Any Time, Post-Baseline-Edaravone Population -On-treatment Period	IP	14.3.4.2
14.3.4.3	Hematology, Categorical Summary of Potentially Clinically Significant Values at by Visit-Safety Population-On-treatment Period	IP	
14.3.4.3.1	Hematology, Categorical Summary of Potentially Clinically Significant Values at by Visit- Edaravone Population-On-treatment Period	IP	14.3.4.3
14.3.4.4	Hematology, Categorical Summary of In/Out of Normal Range Values at by Visit-Safety Population-On-treatment Period	IP	
14.3.4.4.1	Hematology, Categorical Summary of In/Out of Normal Range Values at by Visit-Edaravone Population-On-treatment Period	IP	14.3.4.4
14.3.4.5	Chemistry, Descriptive Summary of Parameters by Visit-Safety Population-On-treatment Period	IP	14.3.4.1
14.3.4.5.1	Chemistry, Descriptive Summary of Parameters by Visit (on-treatment Period)-Edaravone Population-On-treatment Period	IP	14.3.4.1
14.3.4.6	Chemistry, Categorical Summary of Potentially Clinically Significant Values at Any Time, Post-Baseline-Safety Population-On-treatment Period	IP	14.3.4.2
14.3.4.6.1	Chemistry, Categorical Summary of Potentially Clinically Significant Values at Any Time, Post-Baseline-Edaravone Population-On-treatment Period	IP	14.3.4.2
14.3.4.7	Chemistry, Categorical Summary of Potentially Clinically Significant Values at by Visit-Safety Population-On-treatment Period	IP	14.3.4.3



Table Number	Table Title	Validation Method	Shell Number (if repeat)
14.3.4.7.1	Chemistry, Categorical Summary of Potentially Clinically Significant Values at by Visit-Edaravone Population-On-treatment Period	IP	14.3.4.3
14.3.4.8	Chemistry, Categorical Summary of In/Out of Normal Range Values at by Visit-Safety Population-On-treatment Period	IP	14.3.4.4
14.3.4.8.1	Chemistry, Categorical Summary of In/Out of Normal Range Values at by Visit-Edaravone Population-On-treatment Period	IP	14.3.4.4
14.3.4.9	Creatinine, Descriptive Summary by Visit-Safety Population-Patients who have used a MATE/OTC-02 medication during the treatment period-On-treatment Period	IP	14.3.4.1
14.3.4.10	Creatinine, Descriptive Summary by Visit-Safety Population-Patients who did not use a MATE/OTC-02 medication during the treatment period-On-treatment Period	IP	14.3.4.1
14.3.4.11	Lab Parameters, Categorical Summary of Potentially Clinically Significant Categories by Post-Baseline Visit-Safety Population-On-treatment Period	IP	
14.3.4.11.1	Lab Parameters, Categorical Summary of Potentially Clinically Significant Categories by at Any Time, Post-Baseline-Safety Population-On-treatment Period		
14.3.4.11.2	Adverse Events in patients with ALAT or ASAT $\geq 3x$ ULN, Listing-Safety Population-In-trial Period		
14.3.4.11.3	Adverse Events in patients with ALAT or ASAT $\geq 3x$ ULN in combination with esinophils $> 5\%$ , Listing-Safety Population-In-trial Period		
14.3.4.11.4	Lab parameters, listing of patients with ALAT or ASAT $\geq 3x$ ULN in combination with esinophils $> 5\%$ , Listing-Safety Population-In-trial Period		
14.3.4.12	Hematology Lab Parameter Results where at least one value is Abnormal Clinically Significant, Listing-Safety Population-In-trial Period	IP	
14.3.4.13	Chemistry Lab Parameter Results where at least one value is Abnormal Clinically Significant, Listing-Safety Population-In-trial Period	IP	14.3.4.13



Table Number	Table Title	Validation Method	Shell Number (if repeat)
14.3.4.14	Hematology Lab Parameter Results where at least one value is Potentially Clinically Significant, Listing-Safety Population-In-trial Period	IP	14.3.4.12
14.3.4.15	Chemistry Lab Parameter Results where at least one value is Potentially Clinically Significant, Listing-Safety Population-In-trial Period	IP	14.3.4.12
<b>14.3.5</b>	<b>Extent Of Exposure, Dosage Information, And Compliance</b>		
14.3.5.1	Extent of Exposure and mode of Dose administration-Safety Population-In-trial Period	IP	
14.3.5.2	Investigational Medicinal Product Interruptions-Safety Population	IP	
14.3.5.3	Investigational Medicinal Product Interruptions-Safety Population-In-trial Period	IP	
<b>14.3.6</b>	<b>Vital Signs and Physical Examination</b>		
14.3.6.1	Vital Signs, Descriptive Summary of Parameters by Visit-Safety Population-On-treatment Period	IP	14.3.4.1
14.3.6.1.1	Vital Signs, Descriptive Summary of Parameters by Visit-Edaravone Population-On-treatment Period	IP	14.3.4.1
14.3.6.2	Vital Signs, Categorical Summary of Potentially Clinically Significant Categories by Post-Baseline Visit-Safety Population-On-treatment Period	IP	
<b>14.3.7</b>	<b>Other Safety</b>		
14.3.7.1	ECG, Categorical Summary of Overall Interpretation of findings by Visit-Safety Population-On-treatment Period	IP	
14.3.7.1.1	ECG, Categorical Summary of Overall Interpretation of findings by Visit-Safety Population-On-treatment Period	IP	14.3.7.1
14.3.7.2	C-SSRS, Number and Percentage of Patients with a Response of Yes to any of the Suicidal Ideation Items-(timeframe assessed: lifetime and 18 months prior to Screening)-Safety Population-In-trial Period	IP	
14.3.7.3	C-SSRS, Number and Percentage of patients with a Response of Yes to Any Suicidal Behavior and Non-suicidal self-injurious behavior Items (timeframe	IP	



Table Number	Table Title	Validation Method	Shell Number (if repeat)
	assessed: lifetime and 18 months prior to Screening)-Safety Population-In-trial Period		
14.3.7.4	C-SSRS, Number and Percentage of patients with Suicidal ideation, Suicidal Behavior and Self-Injurious Behavior without Suicidal Intent at Any Time Post-Baseline-Safety Population-On-treatment Period	IP	
<b>14.3.8</b>	<b>Concomitant Medication</b>		
14.3.8.1	Prior Medications – Safety Population	IP	
14.3.8.1.1	Prior Medications – Edaravone Population	IP	14.3.8.1
14.3.8.2	Concomitant Medications – Safety Population-In-trial Period	IP	14.3.8.1
14.3.8.2.1	Concomitant Medications – Edaravone Population-in-trial Period	IP	14.3.8.1



Figure Number	Figure Title	Validation Method	Shell Number (if repeat)
14.2.1.1	CAFs, CDF plot-mITT Population-In-trial Period		
14.2.1.2	CAFs, Grouped Histogram-mITT Population-In-trial Period		
14.2.1.3	CAFs, Scatter plot-mITT Population-In-trial Period		
14.2.2.1	Time to PAV/Tracheostomy/Death (including post-withdraw events). Kaplan-Meier-mITT Population-In-trial Period		
14.2.2.4.1	Weibull, Time to PAV/Tracheostomy/Death, Integrated hazard plotted against time-mITT Population-In-trial Period		
14.2.2.4.2	Weibull, Time to PAV/Tracheostomy/Death, log Integrated hazard plotted against time-mITT Population-In-trial Period		
14.2.2.5.1	ALSFRS-R Total Score by Visit (De facto and while alive)-mITT Population-In-trial Period		
14.2.2.5.2	Spaghetti plot of ALSFRS-R Total Scores by timing of last assessment (monthly interval)-mITT Population-In-trial Period		
14.2.2.5.3	Spaghetti plot of patient's estimated regression line for ALSFRS-R total Score by treatment group-mITT Population-In-trial Period		
14.2.2.8.1	% Predicted SVC by Visit (De facto and while alive)-mITT Population-In-trial Period		
14.2.2.8.2	Spaghetti plot of % Predicted SVC by timing of last assessment (monthly interval)-mITT Population-In-trial Period		
14.2.2.8.3	Spaghetti plot of patient's estimated regression line for % Predicted SVC by treatment group-mITT Population-In-trial Period		
14.2.2.9	Scatter plot of baseline ALSFRS-R and % predicted SVC-mITT population		
14.3.1.1	Hematology, Mean by Visit – mITT population		
14.3.1.2	Chemistry, Mean by Visit – mITT population		14.3.1.1
14.3.1.3.1	EDISH-Safety population		
14.3.1.3.2	Spaghetti plot of ALAT and ASAT values for patients with ALAT or ASAT $\geq$ 3xULN-Safety Population-On-treatment Period		
14.3.1.4	Creatinine by Visit -ITT Patients who have used a MATE/OCT-02 medication during the Treatment Period-Safety population		
14.3.1.5	Creatinine by Visit – ITT Patients who did not use a MATE/OCT-02 medication during the treatment period-Safety population		
14.3.1.6	Scatterplot of Change from baseline in Creatinine levels vs Number days since last dose of MATE/OTC-02 medication by treatment Safety subjects who have used a MATE/OCT-02 medication during the Treatment Period		



<b>Listing Number</b>	<b>Listing Title</b>	<b>Validation Method</b>	<b>Shell Number (if repeat)</b>
<b>16.2</b>	<b>Subject Data Listings</b>		
<b>16.2.1</b>	<b>Discontinued Subjects</b>		
16.2.1.1	Subject Disposition and Reasons for Discontinuation from study-ITT Population-In-trial Period	IP	
16.2.1.2	Early Discontinuation From Study Drug-ITT Population-In-trial Period	IP	
16.2.1.3	Screen Failures-Enrolled Population-In-trial Period	IP	
16.2.1.4	Inclusion/Exclusion-Enrolled Population-Screen Failures-In-trial Period	IP	
16.2.1.5	Analysis Population-ITT Population-In-trial Period	IP	
16.2.1.6	Randomization Scheme-ITT Population-In-trial Period	IP	
<b>16.2.2</b>	<b>Protocol Deviations</b>		
16.2.2.1	Protocol Deviations-ITT Population-In-trial Period	IP	
16.2.2.1.1	Site-Level Subject Related Protocol Deviations-ITT Population-In-trial Period	IP	
<b>16.2.4</b>	<b>Demographic Data</b>		
16.2.4.1	Demographic Data-ITT Population-In-trial Period	IP	
16.2.4.2	Medical History-ITT Population-In-trial Period	IP	
16.2.4.3	ALS Characterization-ITT Population-In-trial Period	IP	
16.2.4.4	Tobacco History-ITT Population-In-trial Period	IP	
<b>16.2.5</b>	<b>Compliance and/or Drug Concentration Data</b>		
16.2.5.1	Prior/Concomitant Medication-ITT Population-In-trial Period	IP	
16.1.5.2	Investigation Medicinal Product Administration: Missed Doses, Extra Doses and/or Temporary IMP interruptions-ITT Population	IP	
16.1.5.3	Investigation Medicinal Product Administration: Changes in Mode of IMP administration-ITT Population	IP	
16.1.5.4	Investigation Medicinal Product Administration: Overall Compliance and Extent of exposure-ITT Population	IP	
<b>16.2.6</b>	<b>Individual Efficacy Response Data</b>		



<b>Listing Number</b>	<b>Listing Title</b>	<b>Validation Method</b>	<b>Shell Number (if repeat)</b>
16.2.6.1	CAFs-(Primary Analysis)-mITT Population-In-trial Period	IP	
16.2.6.1.1	CAFs (Sensitivity Analyses without multiple imputation)-mITT Population-In-trial Period	IP	
16.2.6.1.2	CAFs (Sensitivity Analyses -C2R)-mITT Population-In-trial Period	IP	
16.2.6.1.3	CAFs (Sensitivity Analyses -C2R Best)-mITT Population-In-trial Period	IP	
16.2.6.1.4	CAFs (Sensitivity Analyses -C2R Worst)-mITT Population-In-trial Period	IP	
16.2.6.2	ALSFRS-R total Score and Functional Domains-mITT Population-In-trial Period	IP	
16.2.6.3	ALSFRS-R individual Items-mITT Population-In-trial Period	IP	
16.2.6.4	Time to PAV/Tracheostomy/Death-mITT Population-In-trial Period	IP	
16.2.6.5	SVC-mITT population-In-trial Period	IP	
16.2.6.6	ECAS total Score and Subdomains-mITT Population-In-trial Period	IP	
16.2.6.7	ECAS individual Items-mITT Population-In-trial Period	IP	
16.2.6.8	SEIQol-DW Index Score (Subject and Caregiver)-mITT Population-In-trial Period	IP	
16.2.6.9	SEIQol-DW - Subject individual Items-mITT Population-In-trial Period	IP	
16.2.6.10	SEIQol-DW - Caregiver individual Items-mITT Population-In-trial Period	IP	
16.2.6.11	EQ-5D-5L-mITT Population-In-trial Period	IP	
<b>16.2.7</b>	<b>Adverse Event Listings</b>		
16.2.7.1	Adverse Events-ITT population-in-trial Period	IP	
<b>16.2.8</b>	<b>Individual Laboratory Measurements and Other Safety</b>		
16.2.8.1	Hematology -ITT population-in-trial Period	IP	
16.2.8.2	Chemistry-ITT population-in-trial Period	IP	
16.2.8.3	Vital Signs-ITT population-in-trial Period	IP	
16.2.8.4	ECG Findings-ITT population-in-trial Period	IP	

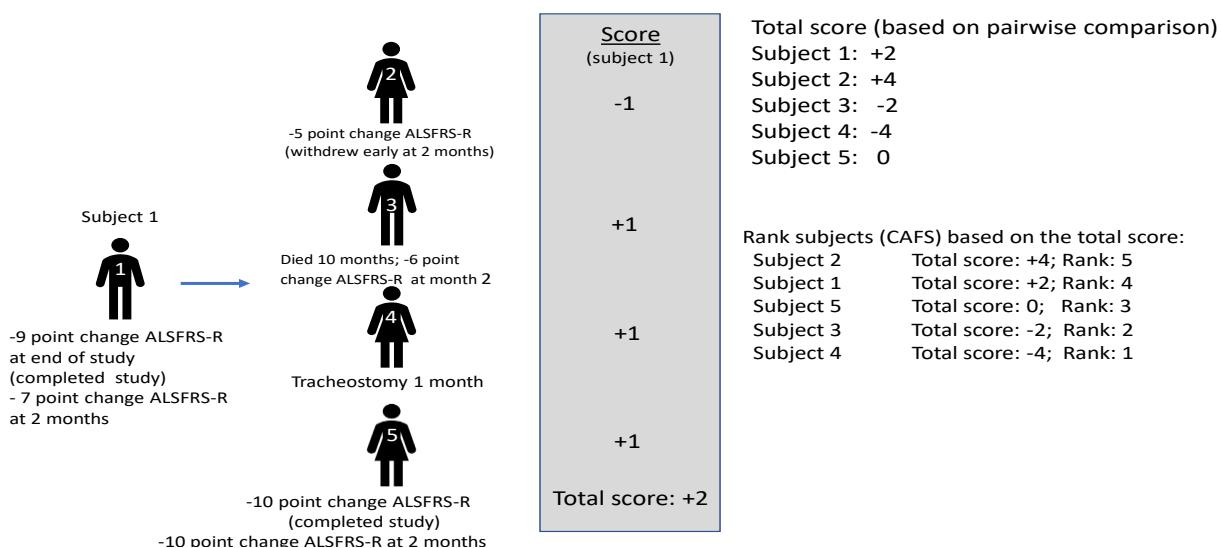


<b>Listing Number</b>	<b>Listing Title</b>	<b>Validation Method</b>	<b>Shell Number (if repeat)</b>
16.2.8.5	Neurological Exam Findings-ITT population-in-trial Period	IP	
16.2.8.6	Physical Exam Findings-ITT population-in-trial Period	IP	
16.2.8.7	C-SSRS-ITT population-in-trial Period	IP	



## 14 APPENDICES

### 14.1 Appendix 1





## 14.2 Appendix 2

### Introduction

The Combined Assessment of Function and Survival (CAFS) was introduced by Berry *et al* (2013) is the primary endpoint. It is a composite endpoint based on change from baseline in ALSFRS-R and survival endpoint (PAV, tracheostomy or death).

The CAFS scoring algorithm is described in Berry *et al* as *If both participants die, the one surviving longer fared better; if only one survives then that participant fared better; and if both participants survive, the one with the smaller decline in ALSFRS-R from baseline fared better. If a participant discontinues early, comparison to each other participant uses time to death if the comparator died before the patient's discontinuation time; otherwise, the comparison is based on the last ALSFRS-R time-point available for both participants.* (The implementation in this trial is to treat the composite survival endpoint as "death" implying, among other things, that there's no follow-up relevant to the CAFS after event.)

The immediate problem is that the above algorithm is ambiguous and could be implemented in different ways based on the same data source. Moreover, the choices to be made seem arbitrary and cannot be said to have little or no impact on the result of the analysis (significance testing).

### Data collected

In the current trial, patients withdrawing from trial are asked whether they would allow follow-up at the time of the planned landmark visit regarding the composite survival endpoint.

For patients accepting this, the observed "history" collected is then a series of ALSFRS-R assessments ending at some intermediate visit followed by the period without observed ALSFRS-R ending with a yes/no on the question on "did patient experience event" (and if "yes" the associated time of event).

At the time writing this, the enquiring whether withdrawals would allow follow-up on the survival endpoint has been implemented haphazardly across trial sites with a presently unknown result.

### The scoring algorithm

In an e-mail correspondence with James Berry, he stressed that the individual subject-to-subject comparisons were to be considered as a prospective comparison. That is, he stressed that if one was to compare, say, a subject with an event at day 200 to a subject lost to follow-up at day 100, and let the subject with an event "loose" the comparison, one would risk introducing a



bias as one couldn't know the destiny of the subject lost to follow-up after day 100. Instead Berry stressed that the comparison was to be performed around the day-100 assessment.

This is important to realize as it then does not allow for comparisons where one would score the trial patients at the landmark visit retrospectively as might otherwise be concluded from Berry *et al* (2013).

As stated above, when implementing the scoring algorithm, a number of arbitrary choices on "data handling" have to be made. For clarity and to enable thorough evaluation of how to construct the primary analysis and sensitivity analyses, the following outlines these choices.

#### *Prospective interpretation and withdrawals*

The prospective interpretation of the subject-to-subject comparisons implies that when we compare two subjects,  $X$  and  $Y$ , the one with the shortest follow-up in essence determines not only the timing but also the "premises" for the subject-to-subject comparison. To see this, consider the following scenarios.

1. Assume that  $X$  has observed ALSFRS-R from  $t = 0, \dots, t_X \leq t_X^d$  where  $t_X^d$  is  $X$ 's event time. Subject  $Y$  has observed ALSFRS-R from  $t = 0, \dots, t_Y < t_X$ . Leaving aside any difficulty in finding matching assessments of the ALSFRS-R assessments, the  $X$  vs.  $Y$  comparison will be based on ALSFRS-R at some time-point  $t = 0, \dots, t_Y$  -- preferably at time  $t_Y$ .

The next scenarios are slightly different:

2A. Assume, as above, that  $X$  has observed ALSFRS-R from  $t = 0, \dots, t_X \leq t_X^d$  where  $t_X^d$  is  $X$ 's event time. Subject  $Y$  has observed ALSFRS-R from  $t = 0, \dots, t_Y < t_X$  but at her withdrawal visit allows the site to follow-up on vital-status. Assume that this follow-up shows that she had an event at  $t_Y^d$ , in which case the comparison will be based on  $t_X^d$  vs.  $t_Y^d$ .

2B. As 2A but now we imagine that the follow-up showed that she didn't have an event but was indeed alive at the time of the planned landmark visit. This would imply that  $t_X^d$  is smaller than the unobserved  $t_Y^d$  and consequently  $X$  would "lose" the comparison.

3. Assume that  $X$  has observed ALSFRS-R from  $t = 0, \dots, t_X$  and subject  $Y$  has observed ALSFRS-R from  $t = 0, \dots, t_Y < t_X$ . In this case, both subjects allowed post-dropout follow-up on vital status which, as it happened, in both cases showed that both were event-free at the time of the planned landmark visit. One therefore has to decide on at which time-point to perform the comparison between ALSFRS-R scores: basically -- do we impute ALSFRS-R scores out to landmark visit?

4. Assume that  $X$  has observed ALSFRS-R from  $t = 0, \dots, t_X$  and subject  $Y$  has observed ALSFRS-R from  $t = 0, \dots, t_Y < t_X$ . In this case, only  $Y$  allowed post-dropout follow-up on vital



status which showed that she was event-free at the time of the planned landmark visit. One therefore has to decide on at which time-point to perform the comparison between ALSFRS-R scores: basically -- do we impute ALSFRS-R scores for  $Y$  up until  $t_X$  (or further)?

-----

[REDACTED]

Because vital status (knowledge of the 3-composite survival endpoint) is an integrated part of the endpoint, ascertaining information on this becomes important. But the fact of the matter is that this information will not be obtained for all subjects (far from it).

Logically it makes little sense to impute ALSFRS-R from the time of withdrawal to  $\min(76w, t^d)$  for withdrawals *allowing* follow-up while at the same time not imputing survival times and ALSFRS-R for withdrawals *not* allowing follow-up. However, Orphazyme has [REDACTED] been advised [REDACTED] not to impute survival-times (for the primary analysis).

[REDACTED] one could impute ALSFRS-R for subjects where vital status would be unknown(able), but this would essentially imply a mixed estimand where these subjects were conditioned on being alive after withdrawal.

#### *Perspectives on "time"*

In Berry *et al* (2013) there is no explicit discussion on "what" constitutes "time": Time is used for two types of subject-subject comparisons. Firstly, when comparing subjects with event(s) the one with the first event "loses" the subject-subject comparison; comparing a subject with an event with onset while the comparator is at risk and under observation *time* per se is used for the comparison. Secondly, when doing a subject-subject comparison where e.g. none of the subjects are observed with events, "time" is used to find matching-in-time assessments of the ALSFRS-R.

As the survival events occur continuously in time and not in connection with visits, obviously event-times are on an actual time scale like *event date minus randomization date +1*.



On the other hand, ALSFRS-R is assessed at visits occurring at discrete/planned time points, say every month, but in reality, may happen haphazardly during the course of the trial. In particular, unscheduled visits occur at time points more or less randomly distributed throughout the trial.

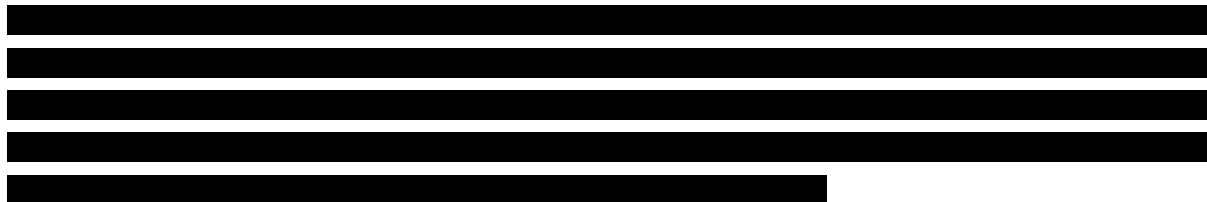
The reason why "unit" of time is relevant is that, when doing subject-subject comparison where none of the subjects are observed with events, finding the assessments to use for the comparison becomes a focus point. When using the actual time of the visits, the likelihood of an exact match to-the-date is small, so some kind of method, with which to find the matching pair, would indeed be warranted. Yet, when using (planned) visits as the unit, where the matching is trivial, other problems have to be dealt with:

1. How to assign assessments to planned visit "slots" including handling of unscheduled visits and dealing with the assessments being "windowed in" to the same visit "slot".
2. How to handle intermittent missing data.

A somewhat distinctive feature of using actual time is that the distinction between withdrawals and those that do not withdraw (actively) from the trial is erased. If one uses actual time, whether the subject's latest ALSFRS-R assessment has anything to do with withdrawing is irrelevant -- all that matters is that observation on the subject ceases at this last assessment.

Another technical issue with using actual time rather than planned visits is that it becomes less obvious how one should impute missing ALSFRS-R data from planned visit after withdrawal conditional on the sequence of (scheduled and unscheduled) visits and their actual timings.

## Analyses

A large rectangular area of the page is completely blacked out with a solid black rectangle, indicating that the content has been redacted for confidentiality.

In list below included in the analyses are all in-trial observed (first) events of the 3-composite survival endpoint along with the specified ALSFRS-R (change from baseline) assessments. The following analyses are therefore planned:

- *Primary Analysis:* CAFS "as is" based on in-trial data (no censoring at time of last dose) using the assessments' actual time and doing (linear) interpolation to get matching ALSFRS-R pairs.



- *Visit windows*: CAFS "as is" based on in-trial data (no censoring at time of last dose) using the in-windowed assessments of ALSFRS-R.
- *Closest match*: CAFS "as is" based on in-trial data (no censoring at time of last dose) using the assessments' actual time and using closest-match to get matching ALSFRS-R pairs.
- *Largest lower*: CAFS "as is" based on in-trial data (no censoring at time of last dose) using the assessments' actual time and using the use largest/latest assessment with actual time  $\leq$  the subject with the shortest follow-up's last assessment to get matching ALSFRS-R pairs.

*Linear interpolation*: Assume that subject  $X$  has observed ALSFRS-R from  $t = 0, \dots, t_X$  and subject  $Y$  has observed ALSFRS-R from  $t = 0, \dots, t_Y < t_X$ . Assume further that  $t_Y$  is not part of subject  $X$ 's sequence of sample time points. Rather, we have that  $X: 0 < t_1 < t_2 < \dots < t_{Y-} < t_Y < t_{Y+} < \dots < t_X$  and corresponding assessments of ALSFRS-R ( $y_{t_h}$  for  $h = 1, \dots$ ). The linear interpolation of  $X$ 's ALSFRS-R data at time  $t_Y$  will then be

$$\widehat{y}_{t_Y} = y_{t_{Y-}} + \frac{y_{t_{Y+}} - y_{t_{Y-}}}{t_{Y+} - t_{Y-}} \cdot (t_Y - t_{Y-}).$$

*Visit windows*: Here the input for the CAFS is the subset of (only) in-windowed ALSFRS-R assessments together with the in-trial observed events of the 3-composite survival endpoint. Matching of ALSFRS-R will be based on visit number/label. Patients with intermittent missing data will have the intermittent missing data imputed by linear interpolation between adjacent observed ALSFRS-R values based on the planned time in months of the visits as "weights".

For a particular patient with intermittent missing data, let the points  $(t_i, y_i)$  for  $i = 0, 1, 2, \dots, L$  with "L" for *landmark* represent the observation schedule (planned time) and associated assessment of the endpoint. The linear interpolation will only use the adjacent observed values for the interpolation:

$t_0$	$t_1$	$t_2$	$t_3$	$t_4$	$t_5 \dots t_L$
$y_0$	$y_1$	$y_2$	$\widehat{y}_3 = y_2 + \frac{y_5 - y_2}{t_5 - t_2} \times (t_3 - t_2)$	$\widehat{y}_4 = y_2 + \frac{y_5 - y_2}{t_5 - t_2} \times (t_4 - t_2)$	$y_5 \dots y_L$

*Closest match*: Assume that subject  $X$  has observed ALSFRS-R from  $t = 0, \dots, t_X$  and subject  $Y$  has observed ALSFRS-R from  $t = 0, \dots, t_Y < t_X$ . Assume further that  $t_Y$  is not part of subject  $X$ 's sequence of sample time points. Rather, we have that  $X: 0 < t_1 < t_2 < \dots < t_{Y-} < t_Y < t_{Y+} < \dots < t_X$  and corresponding assessments of ALSFRS-R ( $y_{t_h}$  for  $h = 1, \dots$ ). The



“closest match” identifies the match to  $(y_{t_Y}, t_Y)$  as  $\min(|t_1 - t_Y|, |t_2 - t_Y|, \dots, |t_X - t_Y|)$  and selects the first of  $X$ ’s assessments in case of a tie.

*Largest lower:* Assume that subject  $X$  has observed ALSFRS-R from  $t = 0, \dots, t_X$  and subject  $Y$  has observed ALSFRS-R from  $t = 0, \dots, t_Y < t_X$ . Assume further that  $t_Y$  is not part of subject  $X$ ’s sequence of sample time points. Rather, we have that  $X: 0 < t_1 < t_2 < \dots < t_{Y-} < t_Y < t_{Y+} < \dots < t_X$  and corresponding assessments of ALSFRS-R ( $y_{t_h}$  for  $h = 1, \dots$ ). The “largest lower” uses  $X$ ’s assessment at time  $t_{Y-}$ .

-----

- *Copy2Reference (C2R):* for subjects with follow-up on survival endpoint after withdrawal, simulation of ALSFRS-R from withdrawal to  $\min(76w, t^d)$  at scheduled time-points beginning with/at the first planned visit after withdrawal. Withdrawn subjects without follow-up / lost to follow-up will be censored at the time of last ALSFRS-R assessment.

While the imputations (see below) *per se* will be based on the in-windowed subset of ALSFRS-R data used for the “Visit Window CAFS” analysis, the C2R-analysis *per se* will include *all* observed ALSFRS-R data augmented with the imputed values using the primary (linear interpolation) CAFS calculation.

The *imputations* will be based on observed in-windowed ALSFRS-R assessments (data used for the “Visit Window CAFS” analysis) and baseline SVC. The imputations will be based on observed placebo subjects (see below).

Two other C2R imputations and associated analyses of CAFS are planned, treating the withdrawn subjects without follow-up on the survival endpoints in two different (extreme) ways

- *C2R worst case:* as detailed above but where the withdrawn subjects without follow-up on the survival endpoints will be treated as if they experienced an event at the time of withdrawal implemented as the time of the last observed ALSFRS-R assessment
- *C2R best case:* as detailed above but where the withdrawn subjects without follow-up on the survival endpoints will be treated as if they did never experience an event during the course of the trial and consequently will have their ALSFRS-R imputed as any other withdrawals up until  $\min(76w, t^d)$ .



### *Details on the Copy2Reference (C2R) imputation*

The imputation approach is based on a pattern mixture modelling (PMM) approach assuming the missing data mechanism is missing at random (MAR) conditional under the observed data under the data selection scheme under the studied estimand.

Under this assumption, the statistical behavior of the missing data can be described by the observed data included in the model, thus, observed values of the endpoint responses at all visits and the observed effects adjusted for in the model.

Consequently, MAR will reflect randomization and allowing follow-up after randomization. Thus, no conditioning of adherence to protocol (except coming in for visits) is implied. Of note, an analysis based on MAR may attribute benefits to patients who withdraw from trial altogether regardless of the reason for discontinuation. The imputation of the monotone missing data will accordingly be performed separately within each of the two randomized treatment.

By performing the imputations in separate groups, the model effectively allows parameters to vary freely amongst the two groups. Moreover, the imputation model will include baseline SVC and previous endpoint values as predictors.

More specifically, let  $(y_{i(k),t}^s)$  and  $(y_{i(k),t}^u)$  for  $i(k) = 1, 2, \dots, N(k)$  and  $t = 0, 1, \dots$  represent the observed scheduled (and in-windowed) assessments of the ALSFRS-R for subject  $i$  from the  $k$ 'th treatment at time  $t$ . The imputations will sequentially loop over each post baseline visit (with missing data) using a monotone regression model as predictor equations:

$$Ey_{i(k),t+1}^s = \mu + \beta SVC_{i(k),0} + \gamma_0 y_{i(k),0}^s + \sum_{g=t-2}^t \gamma_g y_{i(k),g}^s,$$

where the sum includes maximally the two previous visits prior to the  $t$ 'th visit (the one being imputed) and (always) baseline. Here the subjects to have  $y_{i(k),t+1}^s$  imputed are included regardless of treatment and adherence whereas the subjects with observed  $SVC_{i(k),0}$  and  $ALSF_{RS-R}_{i(k),g}$ , for  $g = 0, \dots, t$  (i.e. the subjects used to "inform" the imputations) include only subjects randomised to placebo. SAS code for doing this imputation is in Appendix 3.

### *Details on the MAR-based imputation*

We may perform imputation using the exact same imputation model but within the treatment arms separately. This may be used for analyses under the secondary estimand. (With reference to Appendix 3 the MAR based imputations can be achieved by running almost the same SAS code but for each treatment separately and making sure that the argument *ControlGrp* gets updated in each run accordingly.).



## 14.3 Appendix 3

Appendix 3 contains SAS code and is not included.

## 14.4 Appendix 4

Mapping WorldWide protocol deviation categories to Orphazyme defined categories

Orphazyme Category	WW categories
INCLUSION / EXCLUSION CRITERIA	Eligibility Criteria
WITHDRAWAL CRITERIA	Other: Withdrawal
IMP COMPLIANCE	Investigational Product
DISALLOWED CONCOMITANT MEDICATION	Prohibited Medications
OTHER	Informed Consent Other: Subject Trial ID Card Study Procedures Safety Other: Miscellaneous



If a worldwide category is not listed in the above table, the worldwide category by default will be assigned to Orphazyme category of Other.

## 14.5 Appendix 5

<b>MATE/OCT-02 medications (inhibitors and substrates)</b>
AMANTADINE
AMILORIDE
CIMETIDINE
DOPAMINE
FAMOTIDINE
MEMANTINE
METFORMIN
PINDOLOL
PROCAINAMIDE
RANITIDINE
VARENICLINE
OXALIPLATIN
DOFETELIDE
TRIMETROPRIM
VERAPAMIL
LEVOFLOXACIN
CIPROFLOXACIN
MOXIFLOXACIN
PYRIMETHAMINE
ONDANSETRON
QUINIDINE



## 14.6 Appendix 6

Laboratory Test	CDISC	Unit	PCS LOW	PCS HIGH
<b><i>Haematology / Coagulation</i></b>				
B-haemoglobin	HGB	g/dL	≤ 9.5 (women); ≤ 11.5 (men)	≥ 16.5 (women); ≥ 18.5 (men)
B-erythrocytes (red cell count)	RBC	X 10E12/L	≤ 3.5 (women); ≤ 3.8 (men)	≥ 6.0 (women); ≥ 7.0 (men)
B-haematocrit (packed cell volume)	HCT	V/V	≤ 0.32 (women); ≤ 0.37 (men)	≥ 0.50 (women); ≥ 0.55 (men)
B-total leucocyte (white cell count)	WBC	X 10E9/L	≤ 2.8	≥ 16
B-neutrophils/leucocytes	NEUTLE	%	≤ 20	≥ 85
B-eosinophils/leucocytes	EOSLE	%	-	≥ 10
B-basophils/leucocytes	BASOLE	%	-	≥ 10
B-lymphocytes/leucocytes	LYMLE	%	≤ 10	≥ 75
B-monocytes/leucocytes	MONOLE	%	-	≥ 15
B-thrombocytes (platelet count)	PLAT	X 10E9/L	≤ 75	≥ 600
P-INR (prothrombin ratio)	INR	Ratio	-	≥ 2.0
B-prothrombin time	PT	Sec	-	≥ 18
<b><i>Liver</i></b>				
S-aspartate aminotransferase	AST	IU/L	-	≥ 3 x ULN
S-alanine aminotransferase	ALT	IU/L	-	≥ 3 x ULN
S-bilirubin	BILI	µmol/L	-	≥ 34
S-bilirubin, direct	BILDIR	µmol/L	-	≥ 12
S-bilirubin, indirect	BILIND	µmol/L	-	≥ 22



Laboratory Test	CDISC	Unit	PCS LOW	PCS HIGH
S-alkaline phosphatase	ALP	IU/L	-	$\geq 3 \times \text{ULN}$
S-gamma glutamyl transferase	GGT	IU/L	-	$\geq 200$
S-alpha-glutathione transferase (alpha-GST)	S-GSTAL	$\mu\text{g}/\text{L}$	-	$\geq 20$
<b>Kidney</b>				
S-creatinine	CREAT	$\mu\text{mol}/\text{L}$	-	$\geq 1.5 \times \text{ULN}$
B-urea nitrogen (BUN)	BUN	mmol/L	-	$\geq 11$
<b>Electrolytes</b>				
S-sodium (natrium)	SODIUM	mmol/L	$\leq 125$	$\geq 155$
S-potassium (kalium)	K	mmol/L	$\leq 3.0$	$\geq 6.0$
S-calcium	CA	mmol/L	$\leq 1.8$	$\geq 3.0$
S-chloride	CL	mmol/L	$\leq 90$	$\geq 117$
S-magnesium	MG	mmol/L	$\leq 0.6$	$\geq 1.3$
S-phosphate (phosphorus, (inorganic))	PHOS	mmol/L	$\leq 0.65$	$\geq 1.95$
S-bicarbonate	BICARB	mmol/L	$\leq 12$	$\geq 38$
<b>Endocrine / Metabolic</b>				
B-glucose, non-fasting/unknown	GLUC	mmol/L	$\leq 3.4$	$\geq 9.4$
B-glucose, fasting	GLUC	mmol/L	$\leq 3.0$	$\geq 6.0$
S-glucose, non-fasting/unknown	GLUC	mmol/L	$\leq 3.9$	$\geq 11.1$
S-glucose, fasting	GLUC	mmol/L	$\leq 3.5$	$\geq 7.0$
B-glycosylated haemoglobin, fasting	HBA1C	%	-	$\geq 6.5$
S-prolactin	PROLCTN	mIU/L	-	$\geq 1350$



Laboratory Test	CDISC	Unit	PCS LOW	PCS HIGH
S-protein (total)	PROT	g/L	≤ 45	≥ 95
S-albumin	ALB	g/L	≤ 27	-
<b>Lipids</b>				
S-cholesterol total, non-fasting/unknown	CHOL	mmol/L	-	≥ 7.8
S-cholesterol total, fasting	CHOL	mmol/L	-	≥ 6.2
S-triglycerides, non-fasting/unknown	TRIG	mmol/L	-	≥ 5.65
S-triglycerides, fasting	TRIG	mmol/L	-	≥ 4.2
S-LDL cholesterol, non-fasting/unknown	LDL	mmol/L	-	≥ 5.3
S-LDL cholesterol, fasting	LDL	mmol/L	-	≥ 4.9
S-HDL cholesterol, non-fasting/unknown	HDL	mmol/L	≤ 0.8	-
S-HDL cholesterol, fasting	HDL	mmol/L	≤ 0.9	-
<b>Cardiac / Skeletal / Muscle</b>				
S-creatinine kinase (total)	CK	IU/L	-	≥ 400 (women); ≥ 750 (men)
S-lactate dehydrogenase (total)	LDH	IU/L	-	≥ 750
<b>Infection</b>				
S-C-reactive protein	CRP	mg/L		≥ 25
<b>Urine</b>				
Urinary pH	PH		≤ 4	≥ 9