STATISTICAL ANALYSIS PLAN

VERSION: 2.0

DATE OF PLAN:

24-February-2023

STUDY DRUG:

Reldesemtiv

PROTOCOL NUMBER:

CY 5031

STUDY TITLE:

A Phase 3, Multi-Center, Double-Blind, Randomized, Placebo-Controlled Trial to Evaluate the Efficacy and Safety of Reldesemtiv in Patients with Amyotrophic Lateral Sclerosis (ALS)

BASED ON:

Protocol Amendment 5, 24 February 2023

SPONSOR:

Cytokinetics, Inc.

350 Oyster Point Boulevard, South San Francisco, CA 94080

650-624-3000

This study is being conducted in compliance with good clinical practice, including the archiving of essential documents.

SIGNATURE PAGE

This document has been prepared and/or reviewed by*:

Jenny Wei, PhD, Director, Biostatistics Cytokinetics, Inc.	Signature	Date

This document has been reviewed and accepted by*:

Lisa Meng, PhD, VP, Biometrics, Cytokinetics, Inc.	Signature	Date
Stacy Rudnicki, MD, VP, Clinical Research & Therapeutic Area Lead, Neuromuscular, Cytokinetics, Inc.	Signature	Date
Stuart Kupfer, MD, SVP, Chief Medical Officer Cytokinetics, Inc.	Signature	Date

*See electronic signatures at the end of the document.

TECHNICAL SUMMARY REPORT (TSR)

Name of Sponsor/Company Cytokinetics, Inc.	**Individual Study Table Referring to Part of the Dossier:** **Volume:**	***(For National Authority Use Only)*:**
Name of Finished Product: No generic or trade name assigned	**Page:**	
Name of Active Ingredient: Reldesemtiv (CK-2127107)		
Title of Study: A Phase 3, Multi-Center, Double-blind, Randomized, Placebo-Controlled Trial to Evaluate the Efficacy and Safety of Reldesemtiv in Patients with Amyotrophic Lateral Sclerosis (ALS)		
Investigators: Study Center(s): Patients will be enrolled from approximately 80 clinical trial sites, in the US, Canada, the United Kingdom, Ireland, France, Germany, Italy, Belgium, Portugal, Spain, Sweden, Poland, Switzerland, Australia, and the Netherlands.		
Studied period (years): 2021 to 2023	**Phase of development:** Phase 3	

Objectives and Endpoints:

Objectives	***Endpoint(s)***
Primary	
• To assess the effect of reldesemtiv versus placebo on functional outcomes in ALS	• Change from baseline to Week 24 in ALSFRS-R total score
Secondary	
• To assess the effect of reldesemtiv versus placebo on combined functional and survival outcomes in ALS • To assess the effect of reldesemtiv versus placebo on ventilatory function • To assess the effect of reldesemtiv versus placebo on quality of life • To assess the effect of reldesemtiv versus placebo on handgrip strength	• Combined assessment of change in ALSFRS-R total score, time to dependence on assisted ventilation (use of non-invasive or invasive ventilation for ≥ 22 hours per day for ≥ 10 consecutive days), and survival time up to Week 24 • Change from baseline to Week 24 in the percent predicted FVC • Change from baseline to Week 24 in the ALSAQ-40 total score • Change from baseline to Week 24 in handgrip strength (average of both hands)

Methodology:

This is a Phase 3, double-blind, randomized, placebo-controlled trial of reldesemtiv in patients aged 18 to 80 with ALS. The screening and qualification period for the trial will be no more than 21 days in duration. Approximately 555 eligible ALS patients will be randomized (2:1) to receive the following dose of reldesemtiv or placebo (stratified by riluzole use and edaravone use at baseline) for the first 24 weeks (double-blind placebo-controlled period):

- 300 mg reldesemtiv twice a day for a 600 mg total daily dose (TDD)
- Placebo twice daily

At the end of the 24-week double-blind, placebo-controlled period, patients will transition to the active drug period, where all patients will receive the following dose of reldesemtiv for the next 24 weeks:

- 300 mg reldesemtiv twice a day for a 600 mg TDD for patients who were not down titrated during the 24 weeks of blinded dosing
- 150 mg reldesemtiv twice a day for a 300 mg TDD for patients who were down titrated for any reason during the 24 weeks of blinded dosing

For the entire 48 weeks of dosing, patients and sites shall remain blinded to the randomized treatment assignment patients received during the first 24 weeks.

Study drug should be taken twice daily, morning and afternoon (at least 8 hours apart) and should be taken either with or within the 2-hour period following food.

For patients who complete Week 48 and want to enter the open label extension (OLE), they will not have the follow-up visit; instead, they will enter the OLE on the same day as their Week 48 visit. If the patient does not want to enter the OLE, or if the OLE has not yet been approved at the site, the follow-up visit will take place.

Number of Subjects (planned and analyzed): Approximately 740 patients will be screened in order to randomize/enroll approximately 555 patients, such that approximately 444 evaluable patients complete the 24-week of double-blind placebo-controlled period of the trial.

Diagnosis and main criteria for inclusion:

The key inclusion criteria are below. A full listing of eligibility criteria can be found in protocol Section 5.

- Males or Females between the ages of 18 and 80 years of age, inclusive, at screening.
- Diagnosis of familial or sporadic ALS (defined as meeting the laboratory-supported probable, probable, or definite criteria for ALS according to the World Federation of Neurology El Escorial criteria published in 2000 [Brooks 2000]). Patients who meet the possible criteria are eligible if they have lower motor neuron findings; those who have purely upper motor neuron findings are ineligible.
- First symptom of ALS ≤ 24 months prior to screening. The qualifying first symptoms of ALS are limited to manifestations of weakness in extremity, bulbar, or respiratory muscles. Cramps, fasciculations, or fatigue should not be taken in isolation as a first symptom of ALS.
- ALSFRS-R total score ≤ 44 at screening. Patients with a total score of 45 or higher may be rescreened 60±7 days following the original screening date and be deemed eligible if their ALSFRS-R total score is ≤ 44 or if their score is 2 or more points less than at initial screening. Such patients must continue to meet all other inclusion/exclusion criteria at the time of rescreening.
- Upright FVC ≥ 65.0% of predicted for age, height, sex, and race at screening according to Global Lung Initiative equation (Quanjer 2012).
- Able to perform reproducible pulmonary function tests defined as being able to perform FVC at screening with variability of the 2 highest raw values of less than 10% with a maximum of 5 trials permitted. Screening FVC results must be reviewed and approved by the central review process prior to randomization.
- Must be either on riluzole for ≥ 30 days prior to screening or have not taken it for at least 30 days prior to screening.
- Must have completed at least 2 cycles of edaravone at the time of screening or have not received it for at least 30 days prior to screening.
- Able to swallow whole tablets at the time of screening.
- Clinical laboratory findings within the normal range, or if outside the normal range, not deemed clinically significant by the Investigator, except as specifically indicated as laboratory exclusion.
- Must be either on the combination product (brand name only) of sodium phenylbutyrate and taurursodiol (also known as TUDCA, ursodoxicoltaurine or Tauroursodeoxycholic acid) for ≥51 days or have not received it for at least 30 days prior to screening

Test product, dose and mode of administration:

Reldesemtiv will be administered as 150 mg tablets, at a dose of 300 mg twice a day for a 600 mg TDD. Reldesemtiv tablets will be taken orally.

Patients must interrupt study drug for certain specified laboratory abnormalities. Hepatic related interruptions are permanent discontinuations. For renal related laboratory abnormalities, the interruption continues until the laboratory value in question has returned to the threshold specified in the protocol for re-initiation of treatment. Patients may then resume dosing at 1 tablet twice a day thereafter (150 mg twice daily). Patients also may be down-titrated to 1 tablet twice a day by the Investigator in the case of adverse events felt to be related to study drug. Once patients have been down-titrated to one tablet twice daily for any reason, their dose may not be returned to two tablets twice daily. A physician unblinded to renal laboratory results directs drug interruptions due to renal lab findings.

Duration of treatment:

The treatment period of the study is 48 weeks. Patients randomized to receive 300 mg reldesemtiv twice a day for a 600 mg TDD will receive reldesemtiv for the entire 48-week treatment period; patients randomized to receive placebo twice daily will receive placebo for the first 24 weeks and receive reldesemtiv for the next 24 weeks.

Reference therapy, dose and mode of administration:

Matching placebo will be administered as tablets, at a dose of 2 tablets twice a day. Matching placebo tablets will be taken orally.

In order to maintain the blind to site staff and the study operation team, a small portion of placebo patients will be randomly selected for down-titration to 1 tablet twice a day during the study by the CRO physician unblinded to the renal laboratory results.

Criteria for evaluation (see protocol Section 3):

Efficacy:

The primary efficacy endpoint is change from baseline to Week 24 in ALSFRS-R total score.

The secondary endpoints are as follows:

- Combined assessment of change in ALSFRS-R total score, time to dependence on assisted ventilation (either invasive or non-invasive), and survival time up to Week 24
- Change from baseline to Week 24 in the percent predicted FVC
- Change from baseline to Week 24 in the ALSAQ-40 total score
- Change from baseline to Week 24 in handgrip strength (average of both hands)

Safety:

Patient incidence of adverse events and serious adverse events.

Statistical methods:

Unless otherwise specified, efficacy analyses will be performed on the full analysis set (FAS) by randomized treatment group. The FAS consists of all randomized patients who receive any amount of study drug and have a baseline and at least one post baseline efficacy assessment or have survival status recorded during the 24-week double-blind placebo-controlled period.

The global null hypothesis for the primary endpoint is that there is no treatment difference in the change from baseline to Week 24 in ALSFRS-R total score between patients randomized to placebo and those randomized to reldesemtiv.

The primary analysis will be conducted using a Mixed Model for Repeated Measures (MMRM) with a restricted maximum likelihood method (SAS® PROC MIXED default). The model terms will include treatment group, baseline ALSFRS-R total score, visit, baseline riluzole use, and baseline edaravone use as well as the following interaction terms: baseline ALSFRS-R total score-by-visit and treatment group-by-visit interactions. Other baseline covariates identified by stepwise selection in a regression model will also be included in the primary analysis model if they are key predictor variables for the change from baseline to Week 24 in ALSFRS-R total score based on the blinded data. The regression model used to identify baseline covariates will include pre-study disease progression rate, age and baseline percent predicted FVC as model terms. An unstructured variance-covariance matrix will be used in the primary analysis model.

Prior to applying the MMRM, missing data will be imputed using the multiple imputation procedure (Rubin 1987) for the primary analysis, and the estimates from each imputed dataset will be combined into one set of overall estimates using the SAS® PROC MIANALYZE procedure. The extent of missing data and the pattern of missing data reasons will be tabulated. Least square means (LSM), LSM difference and the corresponding standard errors, 95% confidence intervals (CI) and p-values will be presented. Multiplicity will be addressed using closed testing procedure detailed in Section 8.5.

Safety data will be summarized using descriptive statistics for the safety analysis set that consists of patients who take at least one dose of study drug. All AE terms will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). Primary system organ class (SOC) and preferred term (PT) of the TEAEs will be tabulated by the early-start and delayed-start treatment groups and by the study periods (double-blind placebo-controlled period and active drug period). TEAEs will also be summarized by severity and relationship to study drug. AEs that led to early discontinuation from treatment or trial will be summarized.

TABLE OF CONTENTS

LIST OF TABLES

1. LIST OF ABBREVIATIONS

Abbreviation	Term
AE	Adverse Event
ALP	Alkaline Phosphatase
ALT	Alanine Aminotransferase
ALS	Amyotrophic Lateral Sclerosis
ALSAQ-40	Amyotrophic Lateral Sclerosis Assessment Questionnaire
ALSFRS-R	Amyotrophic Lateral Sclerosis Functional Rating Scale-Revised
ALT	Alanine Aminotransferase
ANCOVA	Analysis of Covariance
AST	Aspartate Aminotransferase
AUC	Area under the ROC Curve
BDI-FS	Beck Depression Inventory (BDI®) - Fast Screen
CI	Confidence Interval
CP	Conditional Power
CYP3A4	Cytochrome P450 Family 3 Subfamily A Member 4
eCRF	Electronic Case Report Form
CRO	Contract Research Organization
CSR	Clinical Study Report
DMC	Data Monitoring Committee
DME	Durable Medical Equipment
EC	Ethics Committee
ECG	Electrocardiogram
EQ-VAS	EuroQol - Visual Analogue Scale
FAS	Full Analysis Set
FVC	Force Vital Capacity
HHD	Hand Held Dynamometry
ICH	International Council for Harmonisation
IRB	Institutional Review Board
LLN	Lower Limit of Normal
LSM	Least Squares Mean
MAR	Missing At Random
MedDRA	Medical Dictionary for Regulatory Activities Terminology

Abbreviation	Term
MiToS	Milano-Torino Staging
MMRM	Mixed Model for Repeated Measures
MNAR	Missing Not At Random
NIV	Non-Invasive Ventilation
OCT1	Organic Cation Transporter 1
OLE	Open Label Extension
PCS	Potential Clinical Significance
PK	Pharmacokinetics
PT	Preferred Term
REB	Research Ethics Board
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SAS	Statistical Analysis System (SAS®)
SOC	System Organ Class
TDD	Total Daily Dose
TEAE	Treatment-Emergent Adverse Event
TESAE	Treatment-Emergent Serious Adverse Event
ULN	Upper Limit of Normal
VAS	Visual Analogue Scale
WHO	World Health Organization

2. INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to provide a technical elaboration of the planned analyses and detailed data displays to be included in the Clinical Study Report (CSR) for CY 5031 study.

This SAP was developed in accordance with ICH E9 and ICH E9 (R1) guideline. All decisions regarding final analysis, as defined in this SAP document, will be made prior to the study database lock. Further study information can be found in the protocol.

3. STUDY OBJECTIVES AND ENDPOINTS

3.1. Study Objectives

3.1.1. Primary Objective

The primary objective of the study is to assess the effect of reldesemtiv versus placebo on functional outcomes in ALS.

3.1.2. Secondary Objective

The secondary objectives of the study are listed as follows:

- To assess the effect of reldesemtiv versus placebo on combined functional and survival outcomes in ALS
- To assess the effect of reldesemtiv versus placebo on ventilatory function
- To assess the effect of reldesemtiv versus placebo on quality of life
- To assess the effect of reldesemtiv versus placebo on handgrip strength

3.2. Study Endpoints

3.2.1. Primary Endpoints

The primary endpoint of the study is change from baseline to Week 24 in ALSFRS-R total score.

3.2.2. Secondary Endpoints

- Combined assessment of change in ALSFRS-R total score, time to dependence on assisted ventilation and survival time up to Week 24. Dependence on assisted ventilation is defined as using non-invasive or invasive ventilation for ≥22 hours per day for ≥10 consecutive days.
- Change from baseline to Week 24 in the percent predicted forced vital capacity (FVC)
- Change from baseline to Week 24 in the ALSAQ-40 total score
- Change from baseline to Week 24 in handgrip strength (average of both hands)

4. STUDY DESIGN

4.1. Summary of Study Design

This is a Phase 3, double-blind, randomized, placebo-controlled trial of reldesemtiv in patients aged 18 to 80 with ALS. The screening and qualification period for the trial will be no more than 21 days in duration. Approximately 555 eligible ALS patients will be randomized (2:1) to receive the following dose of reldesemtiv or placebo (stratified by riluzole use at baseline and edaravone use at baseline) for the first 24 weeks (double-blind placebo-controlled period):

- 300 mg reldesemtiv twice a day for a 600 mg total daily dose (TDD)
- Placebo twice daily

At the end of the 24-week double-blind, placebo-controlled period, patients will transition to the active drug period, where all patients will receive the following dose of reldesemtiv for the next 24 weeks:

- 300 mg reldesemtiv twice a day for a 600 mg TDD for patients who were not down titrated during the 24 weeks of blinded dosing
- 150 mg reldesemtiv twice a day for a 300 mg TDD for patients who were down titrated during the 24 weeks of blinded dosing

For the entire 48 weeks of dosing, patients, sites, Sponsor personnel and their contractors employed to work in any way on the conduct of the trial (with the exception of the unblinded medical monitor) shall remain blinded to the randomized treatment assignment for the first 24 weeks.

Study drug should be taken twice daily, morning and afternoon (at least 8 hours apart) and should be taken either with food or within a 2-hour period following food.

For patients who complete Week 48 and want to enter the open label extension (OLE), they will not have the follow-up visit; instead, they will enter the OLE on the same day as their Week 48 visit. If the patient does not want to enter the OLE, or if the OLE has not yet been approved at the site, the follow-up visit will take place.

4.2. Definition of Study Drugs

This section describes any study drug: investigational product (IP) (ie, reldesemtiv) or placebo intended to be administered to a trial patient according to the protocol.

Arm Name	Active	Placebo
IP/Product Name	Reldesemtiv	Placebo
Type	Drug	Drug
Dose Formulation	Tablet	Tablet
Unit Dose Strength(s)	150 mg	Matching placebo
Dosage Level(s)	2 tablets twice a day for 600 mg TDD	2 matching placebo tablets twice a day
Route of Administration	Oral	Oral
Use	Experimental	Placebo
IMP and NIMP	IMP	IMP

Arm Name	Active	Placebo
Sourcing	Patheon, Inc. Toronto Regional Operations (TRO) 2100 Syntex Court Mississauga, Ontario L5N 7K9 Canada	Patheon, Inc. Toronto Regional Operations (TRO) 2100 Syntex Court Mississauga, Ontario L5N 7K9 Canada
Packaging and Labeling	IP will be provided in bottles. Each bottle will be labeled as required per country requirement	Placebo will be provided in bottles. Each bottle will be labeled as required per country requirement

4.3. Sample Size Considerations

4.3.1. Sample Size Justifications

With a 2:1 randomization ratio to reldesemtiv and placebo groups, respectively, an initial sample size of approximately 555 patients is required to achieve at least 90% power to detect a 1.8 point treatment difference between reldesemtiv and placebo in the change from baseline to Week 24 in ALSFRS-R total score. This calculation is based on a two-sample t-test with two-sided alpha at 0.05 level, a common standard deviation of 5.5 points, and accounting for data missing and early treatment termination.

4.3.2. Sample Size Re-estimation

The sample size will not be re-estimated based on the results of the 1st interim analysis.

The 2nd interim analysis will be triggered 24 weeks after at least one-third of patients required to be enrolled are randomized, and will be based on the data from all patients who are expected to complete 24 weeks of treatment. The sample size may be re-estimated based on the 2nd interim analysis results being reviewed by the DMC. For the primary endpoint, if futility criterion is not met (see Section 5.1 for the futility criterion) and the conditional power under the current trend at the time of the interim look is within the pre-specified promising zone from 0.40 to 0.90, the DMC may recommend increasing the sample size by a pre-specified fixed number of 150 to ensure sufficient power. This method is referred to as the CDL adaptive method and was initially proposed by Chen, DeMets, and Lan (2004) as later extended by Gao, Ware and Mehta (2008) and Mehta and Pocock (2011) (Chen 2004; Gao 2008; Mehta 2011, respectively). It is used to control type-1 error for sample size increase based on unblinded interim data. Within the pre-specified promising zone and a pre-specified fixed increase of the sample size, the CDL method could allow the final analysis to include all data before and after the interim look without inflating the type-1 error.

4.4. Randomization

Eligible ALS patients will be randomized in a 2:1 ratio to receive 300 mg reldesemtiv twice daily or placebo twice daily, respectively, for the first 24 weeks (double-blind, placebo-controlled period). The stratification factors for randomization are baseline riluzole use and baseline edaravone use.

4.5. Clinical Assessments

4.5.1. Efficacy Assessments

Efficacy assessments include Amyotrophic Lateral Sclerosis Functional Rating Scale-Revised (ALSFRS-R), forced vital capacity (FVC), handgrip strength and muscle strength. Additionally, Milano-Torino staging (MiToS) will be derived based on the ALSFRS-R item scores.

4.5.1.1. Amyotrophic Lateral Sclerosis Functional Rating Scale-Revised (ALSFRS-R)

The Amyotrophic Lateral Sclerosis Functional Rating Scale-Revised (ALSFRS-R) is a validated rating instrument for monitoring the progression of disability in patients with ALS (Cedarbaum 1999). The assessment is based on 12 clinical ratings that are categorized in four domains: bulbar (speech, salivation, swallowing), fine motor skills (handwriting, cutting food and handling utensils, dressing hygiene), gross motor skills (turning in bed and adjusting bedclothes, walking, climbing stairs) and respiratory function (dyspnea, orthopnea, respiratory insufficiency). Each clinical rating is assessed using an ordinal scale that ranges from 0 to 4, with a score of 4 representing normal function and lower scores indicating worse function.

The score of each of the four domains (bulbar, fine motor skill, gross motor skill, and respiratory function) is calculated as the sum of the three ratings within each domain and can range from 0 to 12. If a single rating is missing, the score of the corresponding domain will be set to missing.

The ALSFRS-R total score is calculated as the sum of the 12 ratings and can range from 0 to 48 with higher scores indicating less functional impairment. If a single rating is missing, the ALSFRS-R total score will be set to missing.

The score change from baseline will be calculated as the post baseline score minus the baseline score, so that a negative number for change from baseline indicates greater impairment relative to baseline. A positive treatment difference in the change from baseline (reldesemtiv minus placebo) favors reldesemtiv.

4.5.1.2. Percent Predicted Force Vital Capacity (Percent Predicted FVC)

The forced vital capacity (FVC) measures the amount of air a person can forcefully and quickly exhale after inhaling as deeply as possible. FVC measurements will be collected in clinic and remotely. The patient will perform the maneuver 3 times; up to 5 measurements will be taken if the variability of the two highest measurements is ≥10%. Variability is calculated as (highest FVC measurement - 2nd highest FVC measurement) / highest FVC measurement. FVC tracings for this study will be reviewed centrally by independent pulmonologists. Volume time curves will be reviewed for acceptability guided by criteria from the American Thoracic Society and the European Respiratory Society (American Journal of Respiratory and Critical Care Medicine 2019; Vol 200, Number 8). Tracings with variability ≥10% between the 2 highest values are not acceptable. Those tracings that are viewed as acceptable will be ‘evaluable’ and will be included in the analysis. Furthermore, among FVC measurements collected during a visit, the highest result of the evaluable measurements will be used for analysis as long as the pulmonologist has confirmed it is acceptable. FVC measurements collected in clinic and remotely may be combined for analysis if a systematic difference across the entire study population is not observed; otherwise, they will be analyzed separately given that some clinic sites may only collect FVC data remotely during the COVID-19 pandemic due to safety and local health regulations. The

evaluation of whether the systematic difference exists will be performed before the database lock.

The percent predicted FVC represents the FVC measurement as a percent of the predicted values based on sex, age, height, and race. It is calculated as: Percent Predicted FVC = (Trial FVC / Predicted FVC) x 100 (%).

The predicted FVC is calculated using the Quanjer GLI-2012 Regression Equation (Quanjer 2012) that factor the characteristics of sex, age, height and race into the prediction.

For males,

Predicted FVC = exp [–11.2281 + 2.4135 x ln(Height)(cm) + 0.0865 x ln(Age)(yrs)
– 0.1684 x AfrAm – 0.0405 x NEAsia – 0.1177 x SEAsia – 0.0825 x Other + Mspline]

For females,

Predicted FVC = exp [–10.403 + 2.2633 x ln(Height)(cm) + 0.0234 x ln(Age)(yrs)
– 0.1555 x AfrAm – 0.0262 x NEAsia – 0.1516 x SEAsia – 0.0833 x Other + Mspline]

where

ln( ) = natural log transformation

AfrAm = 1 if a patient is African American; otherwise = 0

NEAsia = 1 if a patient is from North East Asia; otherwise = 0

SEAsia = 1 if a patient is from South East Asia; otherwise = 0

Other = 1 if a patient belongs to other ethnic group except for Caucasian, or has mixed race/ethnicity; otherwise = 0

(See https://www.ers-education.org/guidelines/global-lung-function-initiative/faq/what-reference-equations-do-i-apply-for-non-caucasians/ for the definitions of regions used in the Quanjer GLI-2012 Regression Equation.)

Mspline is determined by age and sex as shown in Section 12.4.

Generally, the percent predicted FVC declines over time in patients with ALS. The change from baseline in percent predicted FVC will be calculated as the post baseline value minus the baseline value, so that a negative number for change from baseline indicates respiratory function decline relative to baseline, and a positive treatment difference (reldesemtiv minus placebo) favors reldesemtiv.

4.5.1.3. Milano-Torino Staging (MiToS)

Milano-Torino staging (MiToS) was developed to measure loss of independence in the 4 key domains on ALSFRS (Chio 2015): movement (walking/self-care), swallowing, communicating and breathing. Each MiToS domain score is derived from related ALSFRS-R item scores as shown in Table 1. The sum of the 4 MiToS domain scores has a range from 0 to 4 that corresponds to MiToS Stage 0, Stage 1, Stage 2, Stage 3 and Stage 4, where Stage 0 represents possible functional involvement but no loss of independence on any domain, and Stages 1 to 4 represent number of domains in which independence was lost. In addition, MiToS Stage 5 represents death.

Table 1: MiToS Domain Scores Mapping from ALSFRS-R Item Scores

Domain	ALSFRS-R Items Used for Mapping	Mapping Algorithms
Movement (walking/self-care)	Q8: Walking 4=Normal 3=Early ambulation difficulties 2=Walks with assistance 1= Non-ambulatory functional moment only 0= No purposeful leg movement Q6: Dressing and hygiene 4= Normal function 3= Independent and complete self-care with effort or decreased efficiency 2= Intermittent assistance or substitute methods 1= Needs attendant for self-care 0= Total dependence	If Q8 or Q6 has an item score=0 or 1 then MiToS movement domain score=1; else MiToS movement domain score=0
Swallowing	Q3: Swallowing 4= Normal eating habits 3= Early eating problems; occasional choking 2= Dietary consistency changes 1= Needs supplemental tube feeding 0= NPO (exclusively parenteral or enteral feeding)	If Q3 has an item score=0 or 1 then MiToS swallowing domain score=1; else MiToS swallowing domain score=0
Communicating	Q1: Speech 4= Normal speech processes 3= Detectable speech with disturbances 2= Intelligible with repeating 1= Speech combined with non-vocal communication 0= Loss of useful speech Q4: Handwriting 4= Normal 3= Slow or sloppy; all words are legible 2= Not all words are legible 1= Able to grip pen but unable to write 0= Unable to grip pen	If both Q1 and Q4 have items scores= 0 or 1 then MiToS communicating domain score=1; else MiToS communicating domain score=0.
Breathing	Q10: Dyspnea 4= None 3= Occurs when walking 2= Occurs when one or more of: eating, bathing, dressing 1= Occurs at rest, difficulty breathing when either sitting or lying 0= Significant difficult, considering using mechanical respiratory support Q12: Respiratory insufficiency 4= None 3= Intermittent use of NIPPV 2= Continuous use of NIPPV during the night 1= Continuous use of NIPPV during the night and day 0= Invasive mechanical ventilation by intubation or tracheostomy	If Q10 has an item score=0 or 1 or Q12 has an item score=0, 1 or 2 then MiToS breathing domain score=1; else MiToS breathing domain score=0

4.5.1.4. Handgrip Strength

Handgrip strength will be measured bilaterally by an electronic hand dynamometer with grip extenders to be attached to the front and/or back handle as needed dependent upon patient's hand size. The width of the grip (with no extender, 1 grip extender or 2 grip extenders) will be used consistently throughout the study. Patients will be asked to squeeze the device with the maximum possible force for a short interval to establish the maximum voluntary contraction. The greater of the two maximum handgrip strength attempts for each hand will be used. The secondary endpoint, change from baseline to Week 24 in handgrip strength, will be analyzed using the average grip strength of both hands.

4.5.1.5. Muscle Strength

Patients' muscle strength will be measured bilaterally using hand held dynamometry (HHD). The HHD testing will measure muscle strength of three intrinsic hand muscle groups tested bilaterally: abductor pollicus brevis, first dorsal interosseous, and abductor digiti minimi. Two measurements will be obtained for each muscle tested. If the 2 measurements have a variability > 0.15 then a 3rd test will be performed. The variability is calculated as: (Max Value - Min Value) / Max Value. The maximum result of the 2 or 3 measurements obtained will be used for analysis.

The muscle strength will be recorded as 0 (pounds) on the electronic case report form (eCRF) for the following situations:

- A patient is able to assume the correct position but unable to exert any force.
- A patient is unable to assume the correct position, and the evaluator is able to determine the cause as muscle weakness or contracture.

The muscle strength of each measured body location will be transformed as a percent change from baseline using the equation: [(post baseline value – baseline value) / baseline value] x 100. The transformed muscle strength will be set to missing if the baseline value is zero.

The mega score is a composite score that averages strength across muscle groups. It is calculated as the mean of the non-missing transformed muscle strength scores among the three muscle groups of the hands measured bilaterally by the HHD.

A greater muscle strength mega score indicates less muscle strength loss, and a positive treatment difference (reldesemtiv minus placebo) favors reldesemtiv.

4.5.2. Safety Assessments

Safety assessments include adverse events and serious adverse events, Beck Depression Inventory-Fast Screen (BDI-Fast Screen), electrocardiograms, laboratory assessments, neurological examinations, physical examinations and vital signs.

4.5.2.1. Beck Depression Inventory (BDI)

The Beck Depression Inventory (BDI®) ─ Fast Screen (BDI-FS) is a shortened version of the BDI-II that evaluates key symptoms of depression. It was developed to factor out co-occurring depression symptomatology that may be related to a medical condition (Steer 1999), and concentrate on factors of depression that relate solely to depression. The BDI-FS contains 7 self-

reported items that evaluate major depressive symptoms over the past 2 weeks, including sadness, pessimism, past failure, loss of pleasure, self-dislike, self-criticalness, and suicidal ideations (Beck 2000; Steer 1999). Individual scale items are scored on a 4-point continuum (0=least, 3=most). The total score is calculated as the sum of the scores from the 7 items and can range from 0 to 21. Higher scores indicate greater depressive severity. Patient's depression level will be classified using the BDI total score as shown in Table 2. The BDI total score will be set to missing if an answer to any of the 7 items is missing.

Table 2: Depression Level Classification by BDI Total Score

Depression Level	BDI Total Score
Minimal Depression	0 – 3
Mild Depression	4 – 6
Moderate Depression	7 – 9
Severe Depression	10 – 21

4.5.3. Pharmacokinetic Assessments

Single blood samples will be collected to evaluate plasma concentrations of reldesemtiv and its metabolite, CK-2127106 at Weeks 4, 12 and 24. An optional intensive PK substudy will be conducted at Week 36. Pharmacokinetic (PK) parameters, such as trough concentration (C_{trough}), maximum observed concentration (C_{max}), T_{max}, AUC_{tau}, AUC_{last} and CL/F (reldesemtiv only) and metabolite ratio, if available, will be calculated.

4.5.4. Health Economic Outcome Measures and Quality of Life Assessments

4.5.4.1. EQ-5D-5L and EQ-VAS

The instrument EQ-5D is a standardized measure of health status for clinical and economic appraisal (EuroQol Group 1990), which consists of two parts: a short descriptive system questionnaire (EQ-5D-3L) and a visual analogue scale (EQ-VAS).

EQ-5D-5L is the 5-level version of EQ-5D introduced to improve the instrument's sensitivity and to reduce ceiling effects (EuroQol Group 2009). The descriptive system comprises 5 dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension has 5 levels indicating no problems, slight problems, moderate problems, severe problems or extreme problems. Five responses with a response from each of the 5 dimensions form a 5-digit number that defines a patient's health state profile. A health state can potentially be assigned a summary index score based on societal preference weights (societal perspective) for the health state. The health state preferences often represent national or regional values and can therefore differ between countries/regions. The health state index scores will be calculated using the composite time trade-off (cTTO) method based on the United States valuation of EQ-5D-5L (Pickard 2019) for patients from the United States and for the FAS. The health state index score ranges from less than 0 to 1 with higher scores indicating higher health utility; a score 0 represents death, negative values represent worse than death, and 1 represents full health.

EQ-VAS rates a patient's perceived heath on a vertical visual analogue scale from 0 to 100, where 0 represents the worst imaginable health and 100 represents the best imaginable health. The VAS can be used as a quantitative measure of health outcome that reflects the patient's own judgement.

4.5.4.2. Use of Durable Medical Equipment (DME)

At screening and on Day 1, patients will be asked if they have any of the listed durable medical equipment (DME) (manual wheelchair, power wheelchair, non-invasive ventilation (NIV), gastrostomy tube, augmentative and alternative communication) and if so, the extent of current DME use (has obtained but not using, first use, daily use or dependent) as defined in Table 3, or patient stopped using and the reason of stopping using (as applicable). In the subsequent visits that collect the DME assessments, patients will be asked if any of the listed DME has been prescribed for them and they have agreed to it since the last visit. And if yes, the date of prescription. Patients will also be asked for the extent of current DME use or if they have changed their extent of DME use since the last visit.

Table 3: Definitions of Receipt, First Use, Daily Use and Dependent of the Durable Medical Equipment

	Durable Medical Equipment				
Category	**Manual Wheelchair**	**Power Wheelchair**	**Non-invasive Ventilation**	**Gastrostomy**	**Augmentative and Alternative Communication**
Receipt	Has received (either through insurance or other source)	Has received (either through insurance or other source)	Has received	Has received	Has received (either through insurance or other source)
First use	Use for any reason and in any location	Use for any reason and in any location	Any use outside of being instructed on how to use it – can be day or night use	Any use outside of being instructed on how to use it or flushing it – can be for the purposes of administering medicines, free water, nutrition	Any use for the purposes of communication outside of being instructed on how to use device
Daily use	Daily use (any duration)	Daily use (any duration)	Daily use (can be day or night use) for 4 or more total hrs/24 hrs	Daily use for replacing or supplementing one or more meals per day	Daily use for the purposes of communication
Dependent	Non-ambulatory, wheelchair dependent	Non-ambulatory, wheelchair dependent	Use 22 or more hrs/day	Use for essentially all nutrition, may still take food by mouth for pleasure	Use for all communication

4.5.4.3. Amyotrophic Lateral Sclerosis Assessment Questionnaire (ALSAQ-40)

The Amyotrophic Lateral Sclerosis Assessment Questionnaire (ALSAQ-40) is a 40-item questionnaire that was created specifically to assess health-related quality of life in patients with ALS. ALSAQ-40 includes 40 items for the 5 domains: physical mobility (10 items), activities of daily living and independence (10 items), eating and drinking (3 items), communication (7 items) and emotional reactions (10 items). Based on a two-week recall of experiences patients may have had, each item is rated by frequency of occurrence on a five-point Likert scale (1 = never, 2 = rarely, 3 = sometimes, 4 = often, 5 = always or cannot do at all). For each domain, the sum of scores from all items of a domain is then converted to a summary score from 0 (best health status) to 100 (worst health status) using the equation: (sum of all non-missing item scores / maximum possible sum of all non-missing item scores) x 100. The ALSAQ-40 total score is calculated as the sum of the summary scores from the 5 domains. A lower score corresponds to better health-related quality of life.

4.5.5. Other Research Assessments

Structured patient interviews will be conducted within two weeks before completing the Week 24 visit for a subset of English language-speaking patients (up to 50 patients) enrolled at US centers. The interview will be conducted by phone and will collect information in the patient's words regarding their perceptions of functionality and ALS symptom burden, activities of daily living, and experiences in the trial.

Hospitalizations experienced by a patient following randomization will be determined by the Investigator as being either related to ALS, unrelated to ALS, or indeterminate. Hospitalizations deemed related to ALS cover those that are related to disease progression, hospitalizations that may occur to address an ongoing ALS symptom or to be preventative (such as hospitalization for a PEG / RIG tube), as well as those to address a complication of a treatment being received for ALS. This information will be recorded and entered into the EDC. Hospitalizations that are for social reasons, or those that were planned prior to the patient signing the ICF are not considered SAEs and therefore relationship to ALS will not be recorded.

Biomarkers are objective measures or indicators of normal biological processes, pathological processes or pharmacological responses to a therapeutic intervention. Biomarker development may be useful to identify disease subtypes, guide therapy, detect a response to therapy, better understand the development of an adverse event, or predict disease severity. Where authorized by the applicable IRB/EC/REB approved informed consent, blood samples for biomarker development will be collected pre-AM dose on Day 1 and Week 24. Plasma samples will be stored for future biomarker analyses.

Where authorized by the applicable IRB/EC/REB approved informed consent and authorized by the patient, a blood sample will be obtained at the Day 1 visit and stored for possible future genetic testing related to this study and for research purposes only.

5. PLANNED ANALYSES

5.1. Interim Analyses

Two interim analyses are planned for this trial. The interim analysis methods and the specifications of the data snapshots will be detailed in an interim analysis plan.

Twelve (12) weeks after approximately one-third or more of the planned sample size is randomized, the first interim analysis may be triggered. Its purpose is to determine if there is a lack of effect of reldesemtiv on the ALSFRS-R in CY 5031, a finding that would be inconsistent with the prior observed effect on the change from baseline at 12 weeks in ALSFRS-R total score in the Phase 2 trial, CY 5022. At this interim analysis, the Data Monitoring Committee (DMC) will receive unblinded efficacy data from the independent, third-party biostatistical group supporting them and assess the effect of reldesemtiv on the change from baseline to Week 12 in the ALSFRS-R total score. At this interim assessment, if the treatment difference for change from baseline to Week 12 in ALSFRS-R total score has a one-sided p-value larger than or equal to 0.5 (ie, the estimated treatment difference of the change from baseline at Week 12 in ALSFRS-R total score is favoring placebo), the DMC may recommend stopping the trial due to futility. This futility analysis based on the data of change from baseline to Week 12 in ALSFRS-R will not have any alpha spending. The Sponsor will monitor the blinded aggregate standard deviation of the change from baseline to Week 12 in ALSFRS-R total score periodically and possibly adjust the specified alpha before this interim analysis is conducted.

Twenty-four (24) weeks after at least one-third of the planned sample size is randomized; the second interim analysis will be triggered. The safety analysis will include all data points collected from available patients at the time of the interim data review. The efficacy analysis will include patients who would be expected to complete 24 weeks of the study by the time of data cut off for the second interim data review, i.e., those who complete 24 weeks of the study and those who are expected to complete 24 weeks of the study but early terminate prior to Week 24. Those patients who have not had the opportunity to reach Week 24 by the time of the data cutoff will not be included in the second interim efficacy analysis. The goal of the second interim analysis is to evaluate whether the proposed Phase 3 trial has adequate power to achieve a statistically significant effect on the primary endpoint in the final primary analysis, given the planned enrollment, or if continuing the trial is futile. At this interim analysis, the DMC will receive unblinded efficacy data from the independent, third-party biostatistical group supporting them and assess the effect of reldesemtiv on the primary endpoint. In terms of futility, the DMC may recommend stopping the trial if the conditional power (CP) under the current trend is less than 0.10 for the primary endpoint. The futility boundary is non-binding. The DMC is to make a recommendation to stop the trial using their collective judgment and the totality of evidence available. The DMC may consider the first secondary endpoint in a similar fashion.

If the CP of primary endpoint is within a pre-specified promising zone from 0.40 to 0.90, the DMC may recommend increasing the sample size by a pre-specified fixed number of 150 unless there is any safety concern or other concerns that preclude the recommendation. This method is referred to as the CDL adaptive method and was initially proposed by Chen, DeMets, and Lan, later extended by Gao, Ware and Mehta, and Mehta and Pocock (Chen 2004; Gao 2008; Mehta 2011, respectively). It is used to control type-1 error for sample size increase based on unblinded interim data. Within the pre-specified promising zone and a pre-specified fixed increase of the

sample size, the CDL method could allow the final analysis to include all data before and after the interim look without inflating the type-1 error. This interim analysis will spend a very small 2-sided alpha of 0.0001 for the primary endpoint as well as all individual secondary endpoints. The DMC is also expected to recommend stopping this trial due to superiority if a two-sided p-value is ≤ alpha of 0.0001 for the primary endpoint as well as all individual secondary endpoints. Therefore, the significance level of 2-sided alpha of 0.0499 would be used for the final analyses for the primary endpoint and all individual secondary endpoints to control familywise error rate at overall alpha of 0.05.

5.2. Final Analyses

When all patients complete or discontinue from the study, the final analyses will be conducted.

If the interim analysis result leads to an increase of the sample size, the final primary analyses will be further stratified by the status of being included or excluded from the 2nd interim analysis. As a sensitivity analysis, the adjusted point estimation and confidence interval after the adaption (Lawrence 2003) will also be provided. For the joint rank test on the first secondary endpoint, the ranking will be made separately for the cohort that is included in the 2nd interim analysis and for the subsequent cohort that is not included in order to maintain independence of the results for the cohort included in the interim analysis and the subsequent cohort that is not eligible for inclusion in the interim analysis.

6. GENERAL CONSIDERATIONS FOR DATA ANALYSES AND HANDLING

6.1. General Summary Table and Individual Subject Data Listing Considerations

Unless otherwise specified, analyses will be conducted for the first 24-week double blind placebo-controlled period and the second 24-week active drug period.

Efficacy data will be summarized by randomized treatment group based on the full analysis set (FAS). Safety data will be summarized by treatment received based on the safety analysis set, and PK data will be summarized for patients who receive reldesemtiv based on the pharmacokinetics analysis set.

Descriptive statistics to be presented in a table include count of patients, mean, median, standard deviation, minimum and maximum for continuous variables, and count of patients and the percentage for categorical variables.

For model based analysis, least squares mean, difference of least squares means between treatment groups and the corresponding standard errors, 95% confidence intervals (CIs) and two-sided p-values for the relative statistical inferences will be presented.

For survival analyses, time to event will be summarized by median, 95% CI of median, and 1st and 3rd quartiles using Kaplan-Meier technique. Counts and percentages of events and censored will be presented. Hazard ratio between treatment groups and the corresponding 95% CI and p-value will also be presented.

Listings will include patient ID, demographics, randomized treatment group and other relevant items, and sorted by randomized treatment group, patient ID and date of assessment.

6.2. General Post Text Summary Table and Individual Subject Data Listing Format Considerations

Post text tables and individual subject data listings are prepared according to ICH Guideline E3.

6.3. Data Management

Data will be entered into the clinical database with programmed edit checks and direct visual inspection to ensure integrity. Adverse events, concomitant medications and medical history will be coded programmatically while unique terms will be reviewed visually. Clinical safety laboratory, ECG and PK data and FVC measurements collected from a spirometry at remote visits will be provided in the pre-specified format from external laboratories or research organizations. ALSAQ-40 data collected on CRF and from spirometry will be combined for analysis. Every attempt will be made to obtain measurements to minimize the percentage of missing data.

Critical data will be identified at an early stage of the study to allow sufficient time to have data entered, reviewed, queried and cleaned throughout the study. In preparation for the interim data snapshot, clean CRFs will be soft-locked, which will prevent sites from modifying data on those CRFs.

The final analyses will include data included in the interim analyses, plus additional data that are collected after the interim analyses. The data that are soft locked for the interim analyses will not be modified for the final analyses unless permission is granted from Cytokinetics that a site needs to change the data based on the SOPs. The data management group may break a soft lock to issue a query as needed. Changes for log records such as AE end dates may be updated as appropriate.

6.4. Data Presentation Conventions

The following conventions will be applied to data presentations:

- For continuous variables, mean and median values are formatted to one more decimal place than the measured value. Standard deviation values are formatted to two more decimal places than the measured value. Minimum and maximum values are presented with the same number of decimal places as the measured value.
- For categorical variables, the count and percentage of responses are presented in the form XX (XX.X%).
- Date variables are formatted as DDMMMYYYY for presentation. Time is formatted from 0:00 to 23:59 as HH:MM for presentation.
- P-values, if applicable, will be presented to 4 decimal places. If the p-value is less than 0.0001 then it will be presented as <0.0001. If the rounded result is a value of 1.000, it will be displayed as >0.9999.
- Unless otherwise stated, any statistical tests performed following the second interim analysis (which will be performed at a significance level of 0.0001) will use 2-sided tests at the 0.0499 significance level.

The table and listing shells and table of contents as part of this SAP provide the expected layout and titles of the tables, listings and figures. Any changes to format, layout, titles, numbering, or any other minor deviation will not necessitate a revision to the SAP nor will it be considered a deviation from planned analyses. Only differences in the analysis methods or data handling will require such documentation. The appropriate listings supporting the tables will be included and are not specified in the individual sections throughout the document.

6.5. Analysis Populations

6.5.1. Screen Failures

Screen Failures include patients who sign the informed consent form but are not eligible for study enrollment; therefore, they were not dispensed study drug. Screen failures will not be included in the efficacy and/or safety analyses. However, reasons for screen failures will be summarized. For patients who failed at the first screening and were rescreened, the eligibility at the rescreening visit will be used to determine if the patients are screen failures.

6.5.2. All Randomized Set

All Randomized Set includes patients who are randomized to receive reldesemtiv or placebo.

6.5.3. Safety Analysis Set

Safety analysis set will include all randomized patients who receive any amount of study drug. Patients will be analyzed according to the actual treatment received during the double-blind placebo-controlled period.

6.5.4. Full Analysis Set

Full Analysis Set (FAS) consists of randomized patients who receive any amount of study drug and have a baseline and at least 1 post baseline efficacy assessment or have survival status recorded during the double-blind placebo-controlled period. Patients will be analyzed according to the treatment to which they are assigned at randomization.

6.5.5. Pharmacokinetic Analysis Set

Pharmacokinetic Analysis Set includes all patients who have at least one measurable plasma concentration of reldesemtiv, provided that they have no major protocol violations or deviations that could affect the PK of reldesemtiv.

6.6. Baseline Definition

Unless otherwise specified, baseline is defined as the last available measurement taken prior to administration of the first dose of study drug.

6.7. Derived and Transformed Data

6.7.1. Baseline Age

Age at screening will be collected on the eCRF and used as the baseline age and for all data derivations involving age regardless of age changes during the study.

6.7.2. Study Day

If the date of interest occurs on or after the first dose date, then study day will be calculated as (date of interest – date of first dose) + 1. If the date of interest occurs prior to the first dose date, then study day will be calculated as (date of interest – date of first dose). There is no study day 0.

6.7.3. Change from Baseline

Change from baseline is calculated as (post baseline result – baseline result).

Percent change from baseline is calculated as (change from baseline / baseline result) x 100%.

If either the baseline or the post-baseline result is missing, the change from baseline and/or percentage change from baseline will be set to missing.

6.7.4. Analysis Windows

Since study visits do not always take place exactly as scheduled per protocol, it is necessary to assign the actual observation dates to analysis windows for analysis purposes.

For measurement taken prior to the first dose of study drug, the analysis window is defined as the interval of time prior to the first dose of study drug. Measurements taken on or after the first dose of study drug will be assigned to the next scheduled analysis window.

For data collected at a scheduled post baseline visit, the analysis window will be assigned based on the scheduled study day of the nominal visit as collected on the eCRF.

For unscheduled post baseline visits, early discontinuation visit or follow-up visit, measurements taken on or after the first dose of study drug will be assigned to an analysis window using defined lower and upper bounds for each analysis window. Measurements assigned in an analysis window will have study day greater than or equal to the lower bound but no greater than the upper bound of the analysis window. The lower bound and the upper bound for the analysis windows are defined as the midpoints of the scheduled visits for all assessments (see Section 12.3).

Visits are identified as the nominal visits according to the eCRFs. Each visit will be identified with the visit descriptor (eg, "Week 24"). If a patient has two or more study visits in one analysis window, one record will be flagged as the "analyzed record" for that analysis window.

6.7.5. Multiple Assessments

Once analysis windows are assigned, a patient's individual analysis window could potentially contain more than one visit. Records from all visits, including scheduled, unscheduled and early discontinuation visits could be flagged as the "analyzed record" within the analysis window, although the records from scheduled visit will take priority.

For measurement taken prior to the first dose of study drug, in the event of multiple visits falling within an analysis window, the following rule will be used to determine the "analyzed record":

- The record closest to, but prior to the first dose of study drug should be selected as the baseline "analyzed record." This could include scheduled visits and other screening visits.

For post baseline visits, measurements taken during the 24-week double-blind, placebo-controlled period will be included only in the analysis windows up to Week 24. Measurements taken after patients transition to the active drug period will be included only in the analysis windows after Week 24. Once patients transition to the active drug period, all measurements collected on or after the transition date will not be included in the analysis windows for the first 24 weeks double-blind placebo-controlled period. Furthermore, in the event of multiple visits falling within an analysis window, the following rules will be used in sequence to determine the "analyzed record" for the analysis window:

- If a scheduled visit occurred during the analysis window, then the measurement taken from the scheduled visit will be used.
- If no scheduled visit occurred during the analysis window, the measurement taken closest to the scheduled day will be used as the "analyzed record."
- If no scheduled visit occurred during the analysis window and there is a tie between unscheduled visits in the number of days before and after the scheduled day, measurements from the later visit will be used as the "analyzed record."

For all analyses, only the "analyzed record" within each analysis window will be summarized in a table. If there are other visit records within the analysis window, they will only be included in data listings. In addition, composite scores (eg, ALSFRS-R total score or muscle strength mega score) will be calculated based on the "analyzed record."

6.8. Handling of Missing Data

6.8.1. Missing Efficacy Endpoints

For the primary endpoint, sensitivity analyses with missing data being imputed will be conducted to support the robustness of the analysis result. The base imputation scenario will be based on an ANCOVA model with the MAR-based multiple imputation (Rubin 1987) for all missing ALSFRS-R total scores, and imputation will be performed for each randomized treatment group separately. Imputation based on the distribution of ALSFRS-R total score in the placebo group will also be performed for both reldesemtiv and placebo groups.

Tipping point analyses will be conducted by setting the imputed values for the deaths successively worse only for the reldesemtiv group while maintaining the MAR-based imputed values for the placebo group. Imputation by setting the imputed values successively worse for data missing due to any reasons in the reldesemtiv group will also be performed.

For the first secondary endpoint, sensitivity analyses will be conducted with imputing ALSFRS-R total scores based on the distribution of that in the placebo group. Tipping point analyses will be performed by setting the imputed values successively worse only for the reldesemtiv group while maintaining the MAR-based imputed values for the placebo group.

6.8.2. Missing Start and Stop Dates for Prior and Concomitant Medication

Imputation of missing/partial dates will not be performed. The available year or year and month in a partial date will be used and will be compared to first dosing year, month and day to determine whether to include the medication in the medication history or as a concomitant medication. If the available data do not give sufficient information to classify the medication, the medication will be classified as concomitant medication.

6.8.3. Missing Start and Stop Dates for Adverse Events

For AEs with incomplete date information recorded in the eCRF, the imputation will follow the following algorithm:

For missing AE onset Day and Time:

- If an AE onset Day is missing and the Month of AE onset is known, then the first day of the month of AE onset will be imputed as the AE onset date.
- If AE onset information is not available, then the first dosing date will be imputed as the AE onset date.

For missing AE end Day and Time:

- If the AE end Day is missing and the AE end Month is earlier or later than that of the Follow-Up Visit, then the last day of the AE end month will be the imputed as the AE end date.

- If the AE end Day is missing and the AE end Month is the same as that of the Follow-Up Visit, then the date of the Follow-Up Visit will be the imputed as the AE end date.
- If no AE end information is available, then
  - For patients who discontinued early from study drug, the imputed AE end date will be on the later of the last dosing date + 28 days or the last visit or contact date.
  - For patients who completed the study, the imputed AE end date and time will be the date of the Follow-Up Visit.
  - If the stop date is missing and the event is ongoing, the event will be noted as ongoing" in the stop date column in data listings.

7. STUDY POPULATION

7.1. Subjects Disposition

Count of patients in each analysis set will be presented. The count and percentage of patients who were randomized, who received at least one dose of study drug, who completed the study treatment at each study period, and who prematurely discontinued from the study drug and/or from the study at each study period will be presented by randomized treatment. Reasons for premature discontinuation will also be summarized.

All enrolled patients will be included in the listing of subjects disposition. In addition to patient ID, demographics (age, sex and race) and randomized treatment group, the listing will include screening date, first dose date, last dose date, last contact date, and reasons for discontinuation from the study treatment and/or from the study.

7.2. Screen Failures

Screen failures will be listed and summarized by reasons of screening failure.

7.3. Protocol Deviations

Protocol deviation is any divergence from the protocol that impacts a patient's safety, rights, or welfare or materially reduces the quality or completeness of the data. Patients meeting the protocol deviation criteria will be summarized by randomized treatment group and listed. The listing of protocol deviations will also include severity of deviation (major or minor) and indicate if a protocol deviation is COVID-19 related based on the reported data.

7.4. Demographic and Baseline Characteristics

Demographic and baseline characteristics, including age, sex, race, ethnicity, height, weight, BMI, tobacco and alcohol use, etc., will be summarized by randomized treatment group for the FAS. All enrolled patients will be included in the listing of demographic and baseline characteristics.

7.5. Listing of Subject Inclusion and Exclusion Criteria

A listing of randomized patients who violate the inclusion and exclusion criteria will be provided.

7.6. Medical History and Medical Conditions Present at Entry

Medical history will be summarized by treatment received for the Safety Analysis Set. The count and percentage of patients with each medical history item will be presented by system organ class (SOC) and preferred term (PT). Listing of medical history and medical conditions present at entry will include each medical history item, start date, stop date, system organ class and preferred term.

Medical history of ALS will be summarized by treatment received for the Safety Analysis Set. The count and percentage of patients with or without family history of ALS or a known mutation in an ALS related gene, onset site of ALS, and the descriptive statistics of time since first

symptom onset and time since first diagnosis will be provided. Patient listing containing symptom onset date, and diagnosis date will be provided along with diagnostic time (time between onset of symptoms and diagnosis of ALS).

7.7. Prior Medication History and Medications Present at Entry

7.7.1. Non-Diagnosis Related Prior Medication History

Non-diagnosis related prior medication history will be summarized by treatment received for the Safety Analysis Set. Prior medication history and medication present at entry will be coded using World Health Organization (WHO) Drug Dictionary Enhanced with Herbal, B3, SEP 2020. The count and percentage of patients with each medication history item will be presented by therapeutic class (ATC Class 3) and preferred name. If ATC Class 3 is not available, ATC Class 2 will be used in the summary.

Listing of medication history will include each medication, start date, stop date, therapeutic class and preferred name.

7.7.2. Prior Diagnosis Related Medication History

Prior diagnosis related medication history will be summarized by treatment received for the Safety Analysis Set. The count and percentage of patients who received riluzole and/or edaravone and who received the branded combination product of sodium phenylbutyrate and taurursodiol (also known as TUDCA, ursodoxicoltaurine or Tauroursodeoxycholic acid) will be presented.

7.8. Baseline Physical Examination

Baseline physical examination results will be summarized for the Safety Analysis Set. Counts and percent of patients judged to be normal, abnormal, or not performed will be presented.

7.9. Baseline Vital Signs

Baseline measurements of vital signs, such as heart rate, blood pressure, respiratory rate, weight, BMI and body temperature, will be summarized for the Safety Analysis Set.

7.10. Baseline Laboratory Data

Baseline laboratory data, including hematology, serum chemistry and urinalysis, will be summarized for the Safety Analysis Set. The laboratory parameters to be included in the summary are listed in the protocol Appendix 2.

7.11. Baseline Primary and Secondary Efficacy Evaluations

Baseline measurements of the primary and secondary efficacy analyses will be summarized by randomized treatment group for the FAS with appropriate descriptive statistics.

8. EFFICACY

8.1. General Considerations

Unless otherwise specified, inferential statistical tests will be two-sided and will be performed at alpha levels of 0.05 to declare the significance of effects.

8.2. Testing Statistical Assumptions Including Comparability at Baseline

Assumptions for statistical models will be evaluated graphically. If assumptions are substantially violated, additional analysis methods will be performed. To assess the comparability of treatment groups, demographic and baseline characteristics will be compared between treatments for the FAS using Cochran-Mantel-Haenszel tests for binary categorical variables, van Elteren tests for ordinal categorical measures, or analysis of variance (ANOVA) for continuous variables, stratified by riluzole use and edaravone use at baseline.

8.3. Statement of the Null and Alternate Hypotheses

The null hypothesis (H_0) and alternative hypothesis (H_A) for the primary and the first secondary endpoints are listed below:

For the primary endpoint:

H_0: There is no treatment difference in the change from baseline to Week 24 in ALSFRS-R total score between patients randomized to placebo and those randomized to reldesemtiv during the double-blind, placebo-controlled period in the FAS.

H_A: There is a treatment difference in the change from baseline to Week 24 in ALSFRS-R total score between patients randomized to placebo and those randomized to reldesemtiv during the double-blind, placebo-controlled period in the FAS.

For the first secondary endpoint:

H_0: The Mann-Whitney probability of a more favorable status in either reldesemtiv or placebo group in the FAS is 0.5. The status is quantified by joint ranks based on the combined assessment of change in ALSFRS-R total score, time to dependence on assisted ventilation (either invasive or non-invasive), and survival time up to Week 24.

H_A: The Mann-Whitney probability of a more favorable status in reldesemtiv group in the FAS is greater than 0.5. The status is quantified by joint ranks based on the combined assessment of change in ALSFRS-R total score, time to dependence on assisted ventilation (either invasive or non-invasive), and survival time up to Week 24.

8.4. Subgroup Analyses

8.4.1. Subgroups for Explanation of Potential Heterogeneous Treatment Effect

Subgroup analyses with relatively moderate sample size will be performed to examine the consistency of the observed treatment effect and to gain insight into the effectiveness of reldesemtiv in subpopulations. Analyses of the primary endpoint and the first secondary endpoint will be conducted for the following subgroups:

- Sex (male, female)
- Age group (< median, ≥ median)
- Race (white, non-white)
- BMI at baseline (< 25 kg/m^2, ≥ 25 kg/m^2)
- Riluzole and edaravone use status at baseline (use riluzole, use edaravone, use neither).
- Percent predicted FVC at baseline (< median, ≥ median)
- ALSFRS-R total score at baseline (< median, ≥ median)
- Geographic region (North America, Europe, Australia)
- Anatomic site of disease onset (upper limb, lower limb, limb, bulbar, cognitive, respiratory)
- Time since ALS symptom onset (< 1.5 years, ≥ 1.5 years)
- Pre-study rate of disease progression (1st tertile, 2nd tertile, and 3rd tertile)
- MiToS at baseline (Stage 0, Stage 1, Stage 2, Stage 3 and Stage 4)

For the primary endpoint (ALSFRS-R total score change from baseline to Week 24):

- The MMRM with the subgrouping factors being added to the primary analysis model (see Section 8.6.1) will be used to produce treatment comparisons at each visit for each subgroup. The subgrouping factors include: the "subgroup" listed above as well as "subgroup-by-treatment", "subgroup-by-visit", and "subgroup-by-treatment-by-visit" interactions. If the MMRM fails to converge, an ANCOVA model with change from baseline to Week 24 as response variable and model terms listed as follows will be used to perform the analysis: treatment, baseline riluzole use, baseline edaravone use, baseline ALSFRS-R total score, covariates identified before the database lock (see Section 8.6.1), subgroup and subgroup-by-treatment interaction.

For the first secondary endpoint (combined assessment of change in ALSFRS-R total score, time to dependence on assisted ventilation (either invasive or non-invasive), and survival time up to Week 24):

- The joint rank test subsequent to the MAR imputation of missing ALSFRS-R total scores at Week 24 (see Section 8.7.1) will be applied for all subgroups.

8.4.2. Subgroups Supportive Analyses for the Primary Analyses of Primary Endpoint and First Secondary Endpoint

To explain potential heterogeneity and identify treatment effect modifiers from the baseline characteristics, covariates used to define pre-specified subgroups and these covariates by treatment interaction terms will be included in the primary analysis MMRM model as supportive analysis. Covariates measured as continuous will be introduced to the model as continuous variable. Global test of covariates by treatment interactions will be performed. Stepwise model selection method will be used based on the default stay or entry level of 0.1 to evaluate

significant baseline covariates. The p-value of global test of covariates by treatment group interaction terms as well as the p-values in an ascending order and treatment difference evaluating individual covariates by treatment interactions will be provided from the final multivariate interaction model.

If the primary analysis for the primary endpoint or the first secondary endpoint is significant and the global test of covariates by treatment interactions are having 2-sided p-value less than 0.05, the selected baseline covariate(s) will be considered as treatment effect modifier(s). The corresponding pre-specified subgroup analyses will be considered to explain the potential heterogenous treatment effect.

8.5. Multiple Comparisons and Multiplicity

The null hypothesis for the primary and secondary efficacy endpoints will be tested in a prespecified order using a closed testing procedure. This procedure will maintain the family-wise error rate at two-sided significance level of 0.0499 for all primary and secondary hypotheses tested in a confirmatory sense. The testing steps are described as follows:

Step 1: The null hypothesis for the primary endpoint H_{01} is that there is no treatment difference in the change from baseline to Week 24 in ALSFRS-R total score between patients randomized to placebo and those randomized to reldesemtiv in the FAS. The hypothesis will be tested at the two-sided significance level of 0.0499. If this hypothesis is rejected, testing will proceed to Step 2; otherwise testing will stop.

Step 2: The null hypotheses for the first secondary endpoint H_{02} is that the Mann-Whitney probability of a more favorable status in reldesemtiv or placebo group in the FAS is 0.5. The status is quantified by joint ranks based on the combined assessment of change in ALSFRS-R total score, time to onset of respiratory insufficiency, and survival time up to Week 24. The hypotheses will be tested at the two-sided significance level of 0.05. If this hypothesis is rejected, testing will proceed to Step 3; otherwise testing will stop.

Step 3: The null hypotheses H_{03} is that there is no treatment difference in change from baseline to Week 24 in the percent predicted FVC between patients randomized to placebo and those randomized to reldesemtiv in the FAS. The hypotheses will be tested at the two-sided significance level of 0.05. If this hypothesis is rejected, testing will proceed to Step 4; otherwise testing will stop.

Step 4: The null hypotheses H_{04} is that there is no treatment difference in change from baseline to Week 24 in the ALSAQ-40 total score between patients randomized to placebo and those randomized to reldesemtiv in the FAS. The hypotheses will be tested at the two-sided significance level of 0.05. If this hypothesis is rejected, testing will proceed to Step 5; otherwise testing will stop.

Step 5: The null hypotheses H_{05} is that there is no treatment difference in change from baseline to Week 24 in handgrip strength (average of both hands) between patients randomized to placebo and those randomized to reldesemtiv in the FAS. The hypotheses will be tested at the two-sided significance level of 0.05. No further hypothesis testing will be carried out in this closed testing procedure.

8.6. Analysis of the Primary Efficacy Endpoint

8.6.1. Primary Efficacy Analysis

The primary efficacy endpoint of the trial is the change from baseline to Week 24 in ALSFRS-R total score. The estimand of the primary efficacy endpoint is the difference in the mean changes from baseline to Week 24 in ALSFRS-R total score between patients randomized to placebo and those randomized to reldesemtiv in the FAS.

The analysis for the primary efficacy endpoint will be conducted using a mixed model for repeated measures (MMRM) based on a restricted maximum likelihood method (SAS® PROC MIXED default). The model terms will include treatment group, baseline ALSFRS-R total score, visit, riluzole use at baseline, edaravone use at baseline as well as baseline-by-visit and treatment group-by-visit interactions. In addition, other baseline covariates identified by stepwise selection in a regression model will also be included in the model if they are key predictor variables for the change from baseline to Week 24 in ALSFRS-R total score. The regression model used to identify baseline covariates will include change from baseline to Week 24 in ALSFRS-R total score as response variable and pre-study disease progression rate, age and baseline percent predicted FVC as model terms. The stepwise selection process with significance at 0.25 level for a variable to be entered to the model and with significance at 0.1 level for a variable to remain in the model will be used. The best-fit model according to AIC will be selected. The evaluation will be performed based on the blinded data before the database lock. An unstructured variance-covariance matrix will be used for the model. Observed data collected up to Week 24 in the FAS will be included in the model.

Prior to applying the MMRM, multiple imputation with fifty invocations under the MAR paradigm will be performed to impute missing data (see Section 8.6.1.1 for imputation method).

Least squares mean (LSM), LSM difference and the corresponding standard error, 95% CI and p-value will be presented. The components used to derive main estimator of the primary analysis are shown in Table 4.

Sensitivity analyses will be conducted to support the robustness of conclusions based on the primary analysis results. See Section 8.6.2 for sensitivity analyses.

Table 4: Components of the Main Estimator and Analytical Approach for the Primary Efficacy Endpoint

Estimator	Attribute				Analytical Approach
	Population	Variable	Intercurrent event	Population-level summary	
Main Estimator	FAS	Change from baseline to Week 24 in ALSFRS-R total score	ALSFRS-R total score collected up to Week 24 will be included in the MMRM. Missing data will be imputed using the method described in Section 8.6.1.1.	Difference in variable means between reldesemtiv and placebo groups	MMRM based on a restricted maximum likelihood method (SAS® PROC MIXED default). The model terms will include treatment group, baseline ALSFRS-R total score, visit, riluzole use at baseline, edaravone use at baseline as well as baseline-by-visit and treatment group-by-visit interactions and other baseline covariates identified by stepwise selection.

8.6.1.1. Imputation of the Missing Data for the Primary Efficacy Analysis

The extent of missing data and the pattern of missing data reasons will be tabulated. Missing ALSFRS-R total score will be imputed for the primary efficacy analysis described in Section 8.6.1. For each randomized treatment group, the missing data will be filled sequentially starting at the earliest time point and progressing until the last time point based on a multivariate joint normal imputation model (Schafer 1997). The imputation model will include the following terms: ALSFRS-R total score at previous visits including baseline (in reldesemtiv or placebo group), riluzole use and/or edaravone use at baseline, and other covariates selected using stepwise regression for the primary analysis (see Section 8.6.1) such as pre-study disease progression rate, age and/or baseline percent predicted FVC. Fifty set of complete data will be generated for analysis to produce 50 sets of estimates. Rubin's imputation rules (Rubin 1987) will be used to combine the LSM estimates of the treatment difference, 95% confidence interval and p-value. The example SAS code can be found in Section 12.5.

8.6.2. Sensitivity Analyses of the Primary Efficacy Endpoint

Sensitivity analyses for the primary efficacy endpoint will be conducted based on the FAS and with the common outcome variable and population-level summary. The outcome variable is change from baseline to Week 24 in ALSFRS-R total score, and the population-level summary is the difference of the variable means between reldesemtiv and placebo groups. The sensitivity estimators and the corresponding approach for handling intercurrent events, imputation method and analytical method are shown in Table 5.

In addition, analysis of the primary endpoint will be conducted for the subgroups listed below to evaluate consistency of the observed treatment effect. Same analysis method as described in Section 8.6.1 will be applied to support the analysis result of the primary endpoint.

- Patients with down titration to 150 mg BID at Week 24 (yes, no)
- Patients with COVID-19 related intercurrent events requiring hospitalization (yes, no)
- Patients with COVID-19 related fatal event (yes, no)

Table 5: Sensitivity Estimators of the Primary Efficacy Endpoint

Estimator	Approach for Intercurrent Events	Imputation Method	Analytical Method
Sensitivity Estimator 1	Observed ALSFRS-R total score collected up to Week 24 will be included in the analysis and missing data will not be imputed.	Not applicable.	MMRM described in Section 8.6.1.
Sensitivity Estimator 2	All missing ALSFRS-R total score, regardless of which intercurrent event causes data missing, will be imputed under a MAR paradigm.	Imputation will be performed for each randomized treatment group and will include the following factors in the model: ALSFRS-R total score at previous visits including baseline, riluzole use at baseline, and edaravone use at baseline.	ANCOVA model with ALSFRS-R total score change from baseline to Week 24 as the response variable. The model terms will include treatment group, baseline ALSFRS-R total score, riluzole use status at baseline and edaravone use status at baseline.
Sensitivity Estimator 3	For both reldesemtiv and placebo groups, missing ALSFRS-R total score will be imputed based on the distribution of the placebo group.	For each treatment group, missing ALSFRS-R total scores will be imputed using the parameters from the imputation model of the placebo distribution. The following terms will be included in the imputation model: ALSFRS-R total score at previous visits including baseline, riluzole use at baseline, and edaravone use at baseline.	ANCOVA model with ALSFRS-R total score change from baseline to Week 24 as the response variable. The model terms will include treatment group, baseline ALSFRS-R total score, riluzole use status at baseline and edaravone use status at baseline.
Sensitivity Estimator 4	For reldesemtiv group, tipping point analyses will be conducted by setting the imputed values for the deaths successively worse (starting from 5% worse and with an increment of 5% worse for each of the next step) until a reverse result is obtained. (Ratitch 2013). For reasons of missing other than deaths, missing data will be	For reldesemtiv group, an increment of 1 for the range of shift parameters that reverses the study result will be applied in the multiple imputation. For placebo group, the imputation will be performed with the following factors included in the imputation model: ALSFRS-R total score from previous	ANCOVA model with ALSFRS-R total score change from baseline to Week 24 as the response variable. The model terms will include treatment group, baseline ALSFRS-R total score, riluzole use status at baseline and edaravone use status at baseline.

Estimator	Approach for Intercurrent Events	Imputation Method	Analytical Method
	imputed with MAR-based imputed values. For placebo group, missing ALSFRS-R total score will be imputed with MAR-based imputed values.	visits including baseline, riluzole use at baseline, and edaravone use at baseline.	
Sensitivity Estimator 5	For reldesemtiv group, tipping point analyses will be conducted by setting the imputed values for all missing data, regardless of reasons of missing, successively worse (starting from 5% worse and with an increment of 5% worser for each of the next step) until a reverse result is obtained (Ratitch 2013). For placebo group, missing ALSFRS-R total score will be imputed with MAR-based imputed values.	For reldesemtiv group, an increment of 1 for the range of shift parameters that reverses the study result will be applied in the multiple imputation. For placebo group, the imputation will be performed with the following factors included in the imputation model: ALSFRS-R total score from previous visits including baseline, riluzole use at baseline, and edaravone use at baseline.	ANCOVA model with ALSFRS-R total score change from baseline to Week 24 as the response variable. The model terms will include treatment group, baseline ALSFRS-R total score, riluzole use status at baseline and edaravone use status at baseline.
Sensitivity Estimator 6 (COVID-19 related)	Data collected after the occurrences of COVID-19 related intercurrent events will be excluded.	After excluding the data that were collected after the occurrences of COVID-19 related intercurrent events, imputation will be performed for each randomized treatment group and including the following factors in the imputation model: ALSFRS-R total score from previous visits including baseline, riluzole use at baseline, and edaravone use at baseline.	ANCOVA model with ALSFRS-R total score change from baseline to Week 24 as the response variable. The model terms will include treatment group, baseline ALSFRS-R total score, riluzole use status at baseline and edaravone use status at baseline.

8.7. Analysis of the Secondary Efficacy Endpoints

The secondary endpoint(s) of the trial are:

- Combined assessment of change in ALSFRS-R total score, time to dependence on assisted ventilation (either invasive or non-invasive), and survival time up to Week 24
- Change from baseline to Week 24 in the percent predicted FVC
- Change from baseline to Week 24 in the ALSAQ-40 total score
- Change from baseline to Week 24 in handgrip strength (average of both hands)

The analyses of the secondary endpoints are described in Sections 8.7.1, 8.7.2 and 8.7.3.

8.7.1. Analysis of the First Secondary Efficacy Endpoint

The estimand of the first secondary endpoint (combined assessments of change in ALSFRS-R total score, time to dependence on assisted ventilation (either invasive or non-invasive), and survival time up to Week 24) is the stratified Mann-Whitney probability of a more favorable status in the reldesemtiv group compared to the placebo group in the FAS, based on the joint ranks of the combined assessments collected up to Week 24, and accounting for the intercurrent events of death, dependence on assisted ventilation (either invasive or non-invasive) and decline in ALSFRS-R total score. Dependence on assisted ventilation is defined as patients using invasive or non-invasive ventilation for $\geq$ 22 hours per day for $\geq$ 10 consecutive days.

The joint rank comparison method developed by Berry et al. (Berry 2013) will be applied for the first secondary endpoint. Each patient will first be compared with all others in the analysis population in a pairwise manner on their combined assessments. A score of +1, 0 or -1 will be assigned if one has better, tie or worse outcome than the other, respectively. The worst rank is assigned to patients who die during the trial, followed by ranking on severity of the health status for patients who survive according to their dependence on assisted ventilation and ALSFRS-R total score. The ranking hierarchy is listed as follows:

- Patients who died rank worse than patients who survived.
  - Patients who died earlier rank worse than patients who died later.
- Being dependent upon assisted ventilation is likely to permit patients to survive longer and who would have otherwise died without the intervention. Thus, patients who met the definition of dependence on assisted ventilation (either invasive or non-invasive) rank better than those who died, but worse than those who did not meet the definition of dependence on assisted ventilation (either invasive or non-invasive), regardless of whether the decline in ALSFRS-R total score is smaller or not.
  - Among patients who survived but dependence on assisted ventilation (either invasive or non-invasive), patients with the dependence earlier rank worse.
- Among patients who survived and without dependence on assisted ventilation (either invasive or non-invasive), patients with greater decline in ALSFRS-R total score from baseline rank worse.

  - For patients who discontinued early, comparison is based on the last time-point measurements that are both available for the patients in pairs.

- If it is not possible to assign either patient as having the better outcome, then the score assigned to the pair is 0.

For each patient, the assigned scores (+1, -1 or 0) from the pairwise comparisons with all others will be summed, then all patients in the FAS will be ranked according to their summed scores to create a rank score from 1 for the lowest possible summed score to N (total number of patients in the analysis population) for the highest possible summed score. If ties occur, each of the tied summed scores will be assigned the average of the ranks the tied summed scores would have if there were no ties.

A stratified Wilcoxon test will be applied to compare the standardized joint ranks between reldesemtiv and placebo groups (as ranks divided by pooled group sample size plus one) adjusting for baseline riluzole use and baseline edaravone use. The win probability, win ratio, van Elteren (or stratified Wilcoxon) test statistic, and p-value will be presented.

In the event that mortality rate up to week 24 is higher than expected, the first secondary endpoint may be considered as the primary efficacy endpoint before database lock for the Week 24 primary analysis.

8.7.1.1. Imputation of the Missing Data for the First Secondary Efficacy Analysis

Missing ALSFRS-R total score will be imputed for the first secondary analysis. See Section 8.6.1.1 for the imputation method.

Every effort will be made to collect or derive vital status at Week 24 and Week 48. In the event that missing rate of vital status is higher than expected, the missing vital status will be imputed and sensitivity analyses will be performed.

8.7.2. Sensitivity Analysis of the First Secondary Efficacy Endpoint

Sensitivity analyses for the first secondary efficacy endpoint will be conducted based on the FAS and with the common outcome variable and population-level summary. The outcome variable is the joint-rank based on the combined assessments collected up to Week 24, and the population-level summary is the stratified Mann-Whitney probability of a more favorable status in the reldesemtiv group compared to the placebo group, based on the joint ranks of the combined assessments collected up to Week 24. The sensitivity estimators and the corresponding approach for handling intercurrent events, imputation method and analytical method are shown in Table 6. Likewise, sensitivity analyses will be repeated for the joint rank test with the considerations for patients with COVID-19 related deaths vs. not COVID-19 related deaths.

Table 6: Sensitivity Estimator of the First Secondary Efficacy Endpoint

Estimator	Approach for Intercurrent Events	Imputation Method	Analytical Method
Sensitivity Estimator 7	Observed ALSFRS-R total score collected up to Week 24 will be included in the analysis and missing data will not be imputed.	Not applicable.	Ranking and analysis methods described in Section 8.7.1.
Sensitivity Estimator 8	For both reldesemtiv and placebo groups, missing ALSFRS-R total score due to reasons other than death will be imputed based on the distribution of the placebo group.	For each treatment group, missing ALSFRS-R total scores will be imputed using the parameters from the imputation model of the placebo distribution. The following terms will be included in the imputation model: ALSFRS-R total score from previous visits including baseline, riluzole use at baseline, and edaravone use at baseline.	Ranking and analysis methods described in Section 8.7.1.
Sensitivity Estimator 9	For reldesemtiv group, missing ALSFRS-R total score due to reasons other than death will be imputed successively worse (starting from 5% worse and with an increment of 5% worser for each of the next step) using tipping point approach (Ratitch 2013) until a reverse result is obtained. For placebo group, missing ALSFRS-R total score due to reasons other than death will be imputed with MAR-based imputed values. Missing ALSFRS-R total score due to death will be assigned a worst rank, and patients who died earlier rank worse than patients who died later.	For reldesemtiv group, an increment of 1 for the range of shift parameters that reverses the study result will be applied in the multiple imputation. For placebo group, the imputation will be performed with the following factors included in the imputation model: ALSFRS-R total score from previous visits including baseline, riluzole use at baseline, and edaravone use at baseline.	Ranking and analysis methods described in Section 8.7.1.

Table 6: Sensitivity Estimator of the First Secondary Efficacy Endpoint (Continued)

Estimator	Approach for Intercurrent Events	Imputation Method	Analytical Method
Sensitivity Estimator 10 (COVID-19 related)	Missing ALSFRS-R total score due to COVID-19 related deaths or reasons other than death will be imputed under a MAR paradigm. Missing ALSFRS-R total score due to not COVID-19 related death will be assigned a worst rank, and patients who died earlier rank worse than patients who died later. Ranking from worst to best in the following order will be applied: • Not COVID-19 related death (worst) • COVID-19 related death • With dependence on assisted ventilation (either invasive or non-invasive) • More decline in ALSFRS-R total score • Less decline in ALSFRS-R total score (best)	The imputation will be performed for each randomized treatment group and including the following factors in the model: ALSFRS-R total score from previous visits including baseline, riluzole use at baseline, and edaravone use at baseline.	Ranking and analysis methods described in Section 8.7.1.

8.7.3. Analysis of the Remaining Secondary Efficacy Endpoints

For the secondary endpoint of change from baseline to Week 24 in the percent predicted FVC, analysis will be based on in clinic measurements collected from the Easy One (in clinic) spirometers. Additional analyses of percent predicted FVC that include remote measurements collected from the NuvoAir spirometers will be explored (see Section 8.8).

Analyses for the remaining secondary endpoints will be performed using a MMRM with the following model terms: treatment group, baseline ALSFRS-R total score, visit, riluzole use at baseline, edaravone use at baseline as well as baseline-by-visit and treatment group-by-visit interactions. An unstructured variance-covariance matrix will be specified in the models.

8.8. Analysis of the Exploratory and Additional Efficacy Endpoints

The exploratory endpoints of the trial are:

- Time to the patient being prescribed and patient agrees to any of the following DME items (manual wheelchair, power wheelchair, augmentative and alternative and augmentative communication device, NIV and/or gastrostomy tube),time to first receipt, time to first use, time to daily use, time to dependence and number used of any of the specified DME items from baseline to the Week 24 and Week 48 as well as reasons to receive the above DME
- Changes from baseline to Week 24 and Week 48 in the four subdomain scores of the ALSFRS-R
- Time spent in each MiToS stage and number of patients to transition stages from baseline to Week 24 and Week 48
- Change from baseline to Week 48 in ALSFRS-R total score
- Change from baseline to Week 48 in FVC (liters)
- Change from baseline to Week 48 in grip strength
- Combined assessment of change in ALSFRS-R total score, time to dependence on assisted ventilation (either invasive or non-invasive), and survival time up to Week 48
- Changes from baseline to Week 24 and Week 48 in the mega-score of muscle strength measured by HHD in bilateral first dorsal interosseous muscles, abductor pollicus brevis muscles, and abductor digiti minimi muscles
- Changes from baseline to Week 24 and Week 48 in the EQ-5D-5L
- Changes from baseline to Week 24 and Week 48 in the EQ-Visual Analogue Scale (VAS)
- Time to event for first hospitalization to Week 24 and Week 48

Additional endpoints listed below will be analyzed:

- Time to event for all-cause death from baseline to Week 24 and Week 48
- Time to event for first hospitalization or all-cause death from baseline to Week 24 and Week 48

- Time to event for first ALS related hospitalization or all-cause death from baseline to Week 24 and Week 48
- Time to event for first ALS related hospitalization from baseline to Week 24 and Week 48
- Time to event for first ALS related hospitalization (without hospitalization for a PEG / RIG tube) from baseline to Week 24 and Week 48
- Time to event for first invasive mechanical ventilation by intubation or tracheostomy
- Time to event for first non- invasive mechanical ventilation
- Combined assessment of change in ALSFRS-R total score and survival time up to Week 24 and Week 48 (joint ranks analysis)
- Combined assessment of change in the percent predicted FVC, time to dependence on assisted ventilation (either invasive or non-invasive), and survival time up to Week 24
- Combined assessment of change in the ALSAQ-40 total score, time to dependence on assisted ventilation (either invasive or non-invasive), and survival time up to Week 24
- Change from baseline to Week 24 in FVC (liters)
- Change from baseline to Week 48 in percent predicted FVC
- Slope of change from baseline to Week 24 in percent predicted FVC
- Changes from baseline to Week 24 and Week 48 in forced expiratory volume in the first second
- Change from baseline to Week 24 and Week 48 in percent predicted FVC measured remotely
- Change from baseline to Week 48 in the ALSAQ-40 total score
- Changes from baseline to Week 24 and Week 48 in the subscores of ALSAQ-40
- Time to event for the first intrinsic hand muscle reaches zero from baseline to Week 24 and Week 48
- Time to event for all intrinsic hand muscles reach 0 from baseline to Week 24 and Week 48
- Progression in the degree use of both new and pre-existing DMEs from baseline up to Week 48

The following analyses will also be conducted if a systematic difference between in clinic measurements (from the Easy One spirometers) and remote measurements (from the NuvoAir spirometers) is not observed:

- Change from baseline to Week 24 and Week 48 in percent predicted FVC, with remote measurements collected in the same analysis windows as those used to substitute for the missing in clinic measurements.
- Change from baseline to Week 24 and Week 48 in percent predicted FVC measured in clinic and remotely. In clinic measurements take precedence if both in clinic and remote measurements are collected in the same analysis window.

8.8.1. Analysis of the Exploratory and Additional Efficacy Endpoint

Exploratory analyses will be conducted based on the FAS and by the randomized treatment groups.

Proportional hazard Cox regression models will be applied for the time-to-event endpoints. For patients who had the events, the date of the first event will be used as the event date. For patients who did not have the event, data will be censored at the specified analysis time point (Week 24 or Week 48). Kaplan-Meier method will be used to summarize the time to event, and median, 95% CIs of the median, 1st and 3rd quartiles, and range will be provided.

The joint-rank test as described in Section 8.7.1 will be used to analyze the endpoints of combined assessments up to the specified analysis time points, where the assessments of the endpoints will be used for ranking.

The MMRM as specified for the primary efficacy analysis (see in Section 8.6.1) except for the outcome variable, will be used for the endpoints of change from baseline to a specified time point.

A mixed linear model will be used for the endpoint of slope of change from baseline to a specified time point. The response variable will be the change from baseline to each post baseline visit up to the specified analysis time point, and the model terms will include treatment group, baseline value of the response variable, time from the first dose (a continuous variable), riluzole and edaravone use at baseline, treatment-by-baseline interaction and treatment-by-time interaction. The intercept will be set to zero.

In addition, descriptive statistics will be presented for time spent in each MiToS stage, transition of stages and progression in the degree use of both new and pre-existing DMEs (eg, from daily use progresses to dependence).

8.8.2. Exploratory Analyses Evaluating Effect of Reldesemtiv

To characterize the effect of reldesemtiv, additional comparisons will be conducted based on delayed-start design features.

- The early-start treatment group includes patients who are randomized to receive reldesemtiv. Patients in this group are to receive reldesemtiv during both double-blind, placebo-controlled period and active drug period for a total of 48 weeks.
- The delayed-start treatment group includes patients who are randomized to receive placebo. Patients in this group are to receive placebo during the double-blind, placebo-controlled period for 24 weeks, and then switch to receive reldesemtiv during the active drug period for 24 weeks.

Comparisons listed below will be performed to evaluate treatment differences in 1) change of ALSFRS-R total score from baseline, and 2) combined assessment of change of ALSFRS-R, dependence on assisted ventilation (either invasive or non-invasive), and survival time.

- Compare changes from Week 24 to Week 48 between the early-start and delayed-start treatment groups
- Compare changes from baseline to Week 48 between the early-start and delayed-start treatment groups

- Compare changes from Week 24 to Week 48 versus those from baseline to Week 24 in the delayed-start group

The analysis will include all data collected up to Week 48 (including data from the patients who discontinued the study treatment early but remained in the study and had data being collected). The analysis will also be conducted for patients who were on treatment (ie, excluding data collected after the discontinuation of the study treatment).

8.9. Summary of Reasons for Efficacy Non-Evaluability/Exclusion from Efficacy Analyses

Efficacy analyses will be performed based on the FAS. Reason of excluding patients from the FAS will be summarized or listed.

9. SAFETY AND TOLERABILITY

Safety and tolerability analyses will be based on the Safety Analysis Set. Safety data will be analyzed descriptively and tabulated by the early-start and delayed-start treatment groups for the double-blind placebo-controlled period and active drug period).

9.1. Overall Summary of Tolerability

Overall summary of tolerability will include the following:

- Number of patients treated and patient years of drug exposure (mean, min, max, and median)
- Number of patients with treatment-emergent adverse events (TEAEs)
- Number of patients with treatment-emergent serious adverse events (TESAEs)
- Number of patients with TEAEs leading to permanent dose reduction
- Number of patients with TEAEs leading to premature treatment discontinuation
- Number of patients with TEAEs leading to premature study discontinuation
- Number of Deaths

9.2. Adverse Event Preferred Term and Body/Organ System Summary Tables

9.2.1. Summaries of Adverse Event Incidence Rates for All Subjects

All AE terms will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). TEAEs and TESAEs will be summarized by primary SOC and PT, by PT, and also by severity (Grades 1 to 5) and relationship to study drug (related and not related). For a TEAE reported more than once from a patient, the TEAE will be counted only once in the SOC or PT category using the most severe occurrence and/or closer relationship to the study drug. All AEs will be listed.

9.2.2. Missing and Partial AE Onset Dates

For AEs with incomplete date information recorded on the eCRF, imputation will be performed following the algorithm below:

For missing AE onset Day and Time:

- If an AE onset Day is missing and the Month of AE onset is known, then the first day of the month of AE onset will be imputed as the AE onset date.
- If AE onset information is not available, then the first dosing date will be imputed as the AE onset date.

For missing AE end Day and Time:

- If the AE end Day is missing and the AE end Month is earlier or later than that of the Follow-Up Visit, then the last day of the AE end month will be the imputed as the AE end date.
- If the AE end Day is missing and the AE end Month is the same as that of the Follow-Up Visit, then the date of the Follow-Up Visit will be the imputed as the AE end date.
- If no AE end information is available, then
  - For patients who discontinued early from study drug, the imputed AE end date will be on the later of the last dosing date + 28 days or the last visit or contact date.
  - For patients who completed the study, the imputed AE end date and time will be the date of the Follow-Up Visit.
  - If the stop date is missing and the event is ongoing, the event will be noted as ongoing" in the stop date column in data listings.

9.2.3. Summaries of Adverse Incidence Rates for Serious Adverse Events (SAE), Adverse Event Dropouts, and Death

Serious Adverse Events (SAEs) and AEs leading to premature treatment or study discontinuation and death will be summarized by SOC and PT, and by PT. Listings will be presented for patients who died and/or experienced serious AEs and for patients who discontinued prematurely due to TEAEs.

9.2.4. Summaries of Adverse Events of Special Interest

The following events are considered adverse events of special interest:

- Manifestations of renal impairment and toxicity defined in the protocol Section 7.3, including clinically significant decreases in eGFRCysC and increases of urine protein creatinine ratio
- Clinically significant elevated transaminases and/or bilirubin and manifestations of liver toxicity defined in the protocol Section 7.4 (eg, rash, abdominal pain, nausea and vomiting, fatigue, dark-colored urine, light-colored bowel movements, jaundice, loss of appetite, fever)

Adverse events of special interest recorded on the CRF will be summarized by SOC and PT. Patients with adverse events of special interest and their laboratory values will be listed.

9.3. Total Duration of Therapy, Average Daily Dose, Maximum Daily Dose, Final Daily Dose of Study Medication, and Compliance

Total duration of treatment and exposure of study drug per day will be summarized. All doses taken, date/time of dose, dose reduction (yes/no) and reason of dose reduction will be listed.

9.4. Concomitant and Other Medications

Concomitant medications reported on the eCRF will be summarized. Medications with a start date that is 28 days after the last dose of the study drug will be excluded from the summary. The WHO Drug Dictionary will be used to classify medications by therapeutic class (ATC Class 3) and preferred name. If ATC Class 3 is not available, ATC Class 2 will be used in the summary. Coding will be performed using WHO Drug Dictionary Enhanced with Herbal Dictionary, B3, SEP 2020.

CYP3A4 strong inhibitors should be avoided from 7 days before the start of dosing through the last day of dosing. CYP3A4 strong inducers should be avoided from 14 days before the start of dosing through the last day of dosing. OCT1/OCT2 substrates should be avoided from 7 days before the start of dosing or should be used with caution during the study. Patients who take CYP3A4 strong inhibitors, CYP3A4 strong inducers, or OCT1/OCT2 substrates prior or during the study treatment will be listed.

Concomitant medications taken throughout the study will be listed. The listing will include concomitant medication, start/stop date, study day, dose / unit / route / frequency, indication and purpose of concomitant medications.

9.4.1. Missing and Partial Concomitant and Other Medication Start and Stop Dates

Incomplete concomitant medication start and/or stop dates will be imputed as follows. If either the start of medication date is completely missing or the month and/or year of the start date is missing, the start date will be set to the first dose date. If only the day is missing, the day will be set to the first day of the month except if the start month is the same as the first dose month. If the latter is true, the start day will be set to the first dose day. No imputations will be applied to the stop date.

9.5. Routine Laboratory Data

Clinical chemistry, hematology and urinalysis laboratory measurements and value changes from baseline at each laboratory blood sample collection time point will be summarized. The count and percentage of patients who had normal or missing laboratory values at baseline and abnormal laboratory values post baseline will be presented. The lower limit of normal (LLN) and upper limit of normal (ULN) provided by the laboratories will be used as the criteria to determine abnormality. For each parameter, the denominator of the percentage will include patients with normal or missing assessments at baseline, and with at least one assessment post baseline. The numerator of the percentage will include patients who had at least one abnormal assessment post baseline among the patients that were counted in the denominator. Assessment collected at unscheduled visits or the Follow-up Visit will be included in the summary.

Shift of clinical laboratory results from baseline severity to the maximum post baseline severity will be presented for selected laboratory parameters. Clinical laboratory values will be also listed. Values outside the Laboratory's normal ranges and potentially clinically significant abnormal values will be flagged.

Estimated glomerular filtration rate (eGFR) will be calculated using the CKD-EPI cystatin equation (Inker 2012) and summarized as change from baseline, percentage change from

baseline, by categories (15 to <30, 30 to < 45, 45 to < 60, 60 to <90 and ≥90, as well as ≥10%, ≥25% and >= 50% decline from baseline) and by shift in categories.

Liver function test results will be summarized as change from baseline and as count and percentage of patients with normal baseline and abnormal post-baseline results in Alanine Aminotransferase (ALT), Aspartate Aminotransferase (AST), Alkaline Phosphatase (ALP) and bilirubin, with the following categories:

- ALT > 3xULN, > 5xULN, > 8xULN
- AST > 3xULN, > 5xULN, > 8xULN
- ALT and/or AST > 3xULN, > 5xULN, > 8xULN
- ALT and/or AST > 3xULN and total bilirubin > 2xULN and ALP < 2xULN
- ALT and/or AST > 3xULN and total bilirubin > 2xULN
- Bilirubin (total, indirect or direct) > 2xULN, >3xULN
- ALT or AST > 3 x ULN with symptoms including nausea, vomiting, anorexia, abdominal pain, fatigue, rash, dark-colored urine, light-colored bowel movements, jaundice, or fever

In addition, time to ALT, AST or bilirubin elevation since the first dose will be summarized by treatment groups.

9.6. Vital Signs

Blood pressure, heart rate, respiratory rate, body weight and oral temperature will be collected at the time points specified in the protocol. For each assessment time point, absolute values and changes from baseline of the vital signs will be summarized.

The count and percentage of patients who had normal or missing values at baseline and Potentially Clinically Significant (PCS) values post baseline will be presented. The PCS criteria are specified in Table 7. For each parameter, the denominator of the percentage will include patients with non-PCS or missing assessments at baseline, and with at least one assessment post baseline. The numerator of the percentage will include patients who had at least one PCS assessment post baseline among the patients that were counted in the denominator. Assessment collected during unscheduled visits or the Follow-up Visit will be included in the summary.

All collected vital signs and assessment time will be listed, and values outside the normal ranges will be flagged.

Table 7: Criteria of Potentially Clinically Significant (PCS) Vital Signs

Vital Sign Parameter	Flag	Criteria
Systolic Blood Pressure (mmHg)	High	≥160 mmHg
	Low	≤80 mmHg
Diastolic Blood Pressure (mmHg)	High	≥100 mmHg
	Low	≤ 50 mmHg
Respiration Rate (Breaths per minute)	High	>18 bpm
Respiration Rate (Breaths per minute) (Cont'd)	Low	<8 bpm
Pulse (bpm)	High	≥120 bpm

Table 7: Criteria of Potentially Clinically Significant (PCS) Vital Signs (Continued)

Vital Sign Parameter	Flag	Criteria
	Low	≤50 bpm
Weight (kg)	Clinically Significant	≥5% reduction from baseline
Temperature (°C)	High	≥ 38 °C
	Low	< 35 °C

9.7. Electrocardiogram

A 12-lead ECG will be obtained at the time points specified in the protocol using an ECG machine that automatically calculates the heart rate and measures PR, QRS, QT, and QTc intervals. Descriptive statistics of absolute value and change from baseline of ECG parameters (PR interval, RR interval, QRS duration, QT interval, QTc interval [Bazett's and Fridericia's], Ventricular Heart Rate) at each assessment time point will be presented.

The count and percentage of patients who had normal or missing values at baseline and PCS values post baseline will be presented. ECG parameters are regarded as PCS if the value meets the criterion shown in Table 8. For each parameter, the denominator of the percentage will include patients with non-PCS or missing assessments at baseline, and with at least one assessment post baseline. The numerator of the percentage will include patients who had at least one PCS assessment post baseline among the patients that were counted in the denominator. Assessment collected at an unscheduled visits or the Follow-up Visit will also be included in the summary.

All collected 12-lead ECG parameters, assessment time, ECG findings and interpretations will be listed. PCS ECG values will be flagged.

Table 8: Criteria of Abnormal and PCS ECG Values

ECG Variable	Units	Upper Limit of Normal	PCS High Values
QRS Interval	msec	80	≥120
PR Interval	msec	200	≥240
QTcB Interval	msec	Males: 450 Females: 460	>500
QTcF Interval	msec	Males: 450 Females: 460	>500
QT Interval	msec	—	> 500
Ventricular Heart Rate	bpm	—	≤50 >100

9.8. Physical Examination

Reported abnormal and clinically significant findings in physical examinations will be summarized as Medical History or Adverse Events.

9.9. Beck Depression Inventory Fast Screen Version

The BDI-FS will be assessed at the time points specified in the protocol. Descriptive statistics of the BDI total score will be provided for each assessment time point. The actual assessment date and all reported items will also be listed.

9.10. Study Termination Status

Count and percentage of patients who completed or prematurely discontinued planned study dosing will be presented. Reasons of prematurely discontinuation will be tabulated. For patients who discontinued prematurely, the date and reason of premature discontinuation will be listed.

10. PHARMACOKINETICS

Plasma concentrations of reldesemtiv and its metabolite CK-2127106 at each collection time point will be summarized using descriptive statistics including mean, standard deviation, geometric mean, coefficient of variation, median, and range. Mean plasma concentrations over time will be graphically displayed. PK Parameters, such as C_{trough}, C_{max}, T_{max}, AUC_{tau}, AUC_{last} and CL/F (reldesemtiv only) and metabolite ratio, if available, will be summarized. Population PK modeling based on the plasma concentrations of reldesemtiv from this study and previous studies will be conducted; results will be shown in a separate report.

11. REFERENCES

Beck, A. T., Steer, R. A. and Brown, G. K. (2000). Bdi-fast screen for medical patients manual. San Antonio, TX, Psychological Corporation.

Berry, J. D., Miller, R., Moore, D. H., Cudkowicz, M. E., Van Den Berg, L. H., Kerr, D. A., et al. (2013). "The combined assessment of function and survival (cafs): A new endpoint for als clinical trials." Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration 14(3): 162-168.

Brooks, B. R., Miller, R. G., Swash, M. and Munsat, T. L. (2000). "El escorial revisited: Revised criteria for the diagnosis of amyotrophic lateral sclerosis." Amyotrophic Lateral Sclerosis and Other Motor Neuron Disorders 1(5): 293-299.

Cedarbaum, J. M., Stambler, N., Malta, E., Fuller, C., Hilt, D., Thurmond, B., et al. (1999). "The alsfrs-r: A revised als functional rating scale that incorporates assessments of respiratory function." Journal of the Neurological Sciences 169(1): 13-21.

Chen, Y. H. J., DeMets, D. L. and Gordon Lan, K. K. (2004). "Increasing the sample size when the unblinded interim result is promising." Statistics in Medicine 23(7): 1023-1038.

Chio, A., Hammond, E. R., Mora, G., Bonito, V. and Filippini, G. (2015). "Development and evaluation of a clinical staging system for amyotrophic lateral sclerosis." Journal of Neurology, Neurosurgery & Psychiatry 86(1): 38-44.

EuroQol Group (1990). "Euroqol - a new facility for the measurement of health-related quality of life." Health Policy 16(3): 199-208.

EuroQol Group. (2009). "Eq-5d-5l." from https://euroqol.org/eq-5d-instruments/eq-5d-5l-about/, Updated Date Accessed Date.

Gao, P., Ware, J. H. and Mehta, C. (2008). "Sample size re-estimation for adaptive sequential design in clinical trials." Journal of biopharmaceutical statistics 18: 1184-1196.

Inker, L. A., Schmid, C. H., Tighiouart, H., Eckfeldt, J. H., Feldman, H. I., Greene, T., et al. (2012). "Estimating glomerular filtration rate from serum creatinine and cystatin c." New England Journal of Medicine 367(1): 20-29.

Lawrence, J. and Hung, H. M. J. (2003). "Estimation and confidence intervals after adjusting the maximum information." Biometrical Journal 45(2): 143-152.

Mehta, C. R. and Pocock, S. J. (2011). "Adaptive increase in sample size when interim results are promising: A practical guide with examples." Statistics in Medicine 30(28): 3267-3284.

Pickard, A. S., Law, E. H., Jiang, R., Pullenayegum, E., Shaw, J. W., Xie, F., et al. (2019). "United states valuation of eq-5d-5l health states using an international protocol." Value in Health 22(8): 931-941.

Quanjer, P. H., Stanojevic, S., Cole, T. J., Baur, X., Hall, G. L., Culver, B. H., et al. (2012). "Multi-ethnic reference values for spirometry for the 3–95-yr age range: The global lung function 2012 equations." European Respiratory Journal 40(6): 1324-1343.

Ratitch, B., O'Kelly, M. and Tosiello, R. (2013). "Missing data in clinical trials: From clinical assumptions to statistical analysis using pattern mixture models." Pharmaceutical Statistics 12(6): 337-347.

Rubin, D. B. (1987). Multiple imputation for nonresponse in surveys. New York, John Wiley & Sons, Inc.

Schafer, J. L. (1997). Analysis of incomplete multivariate data. New York, Chapman and Hall/CRC.

Steer, R. A., Cavalieri, T. A., Leonard, D. M. and Beck, A. T. (1999). "Use of the beck depression inventory for primary care to screen for major depression disorders." General Hospital Psychiatry 21(2): 106-111.

12. APPENDIX

12.1. Table of Contents for Data Display Specifications

Table of contents for data display specifications will be provided in a separate document.

12.2. Data Display Specifications

Data display specifications will be provided in a separate document.

12.3. Analysis Windows

Measurements collected during the 24-week double-blind placebo-controlled period will be included only in the analysis windows up to Week 24; measurements collected after patients transition to the active drug period will be included only in the analysis windows after Week 24.

Table 9: Analysis Windows for ALSFRS-R, FVC (Remote Visits) and DME

Visit	Scheduled Day	Lower Bound	Upper Bound
Screening[a]	<1	<1	<1
Day 1	1	1	1
Week 4	29	2	42
Week 8	57	43	70
Week 12	85	71	98
Week 16	113	99	126
Week 20	141	127	154
Week 24	169	155	182
Week 28	197	183	210
Week 32	225	211	238
Week 36	253	239	266
Week 40	281	267	294
Week 44	309	295	322
Week 48	337	323	350
Follow-up	365	351	>351

[a] Not applicable for FVC (Remote Visits)

Table 10: Analysis Windows for FVC (in Clinic), Handgrip and Muscle Strength, BDI-FS, Vital Signs, Weight and BMI

Visit	Scheduled Day	Lower Bound	Upper Bound
Screening	<1	<1	<1
Day 1	1	1	1
Week 4	29	2	56
Week 12	85	57	126
Week 24	169	127	210
Week 36	253	211	294
Week 48	337	295	350
Follow-up	365	351	>351

Table 11: Analysis Windows for EQ-5D-5L, EQ-VAS and ALSAQ-40

Visit	Scheduled Day	Lower Bound	Upper Bound
Screening	<1	<1	<1
Day 1	1	1	1
Week 12	85	2	126
Week 24	169	127	210
Week 36	253	211	294
Week 48	337	295	350
Follow-up	365	351	>351

Table 12: Analysis Windows for Neurological Exam

Visit	Scheduled Day	Lower Bound	Upper Bound
Screening	<1	<1	<1
Week 24	169	1	266
Follow-up	365	267	>267

Table 13: Analysis Windows for ECG

Visit	Scheduled Day	Lower Bound	Upper Bound
Screening	<1	<1	<1
Day 1	1	1	1
Week 24	169	2	252
Week 48	337	253	350
Follow-up	365	351	>351

Table 14: Analysis Windows for PK Concentration

Visit	Scheduled Day	Lower Bound	Upper Bound
Week 4	29	2	49
Week 12	85	50	126
Week 24	169	127	>210
Week 36	253	211	>211

Table 15: Analysis Windows for Clinical Safety Laboratory Measurements

Visit	Scheduled Day	Lower Bound	Upper Bound
Screening	<1	<1	<1
Day 1	1	1	1
Week 2	15	2	21
Week 4	29	22	42
Week 8	57	43	70
Week 12	85	71	98
Week 16	113	99	126
Week 20	141	127	154
Week 24	169	155	175
Week 26	183	176	189
Week 28	197	190	210
Week 32	225	211	238
Week 36	253	239	266
Week 40	281	267	294
Week 44	309	295	322
Week 48	337	323	350
Follow-up	365	351	>351

12.4. Mspline Value for Calculation of Predicted FVC

Table 16: Mspline Value

Age (yrs)	Male	Female	Age (yrs)	Male	Female
18	0.147849088	0.135507529	50	-0.006647706	0.028303776
19	0.160315109	0.140628299	51	-0.014542406	0.020206649
20	0.167159394	0.143647414	52	-0.022604908	0.011905392
21	0.169722148	0.145330204	53	-0.030845244	0.00339698
22	0.16923214	0.146039097	54	-0.039274925	-0.005323975
23	0.166543254	0.145870185	55	-0.04790158	-0.014258162
24	0.162315681	0.144942845	56	-0.056735557	-0.023411261
25	0.157092735	0.143626724	57	-0.065782826	-0.032797239
26	0.151293448	0.142211366	58	-0.07504859	-0.042417299
27	0.145106883	0.140771596	59	-0.084531216	-0.052260185
28	0.138681964	0.13919976	60	-0.094214527	-0.062313185
29	0.132266867	0.137441293	61	-0.104078877	-0.072556666
30	0.125876212	0.135494281	62	-0.114093666	-0.082975133
31	0.119459845	0.133330256	63	-0.124218952	-0.093558031
32	0.113049003	0.130878076	64	-0.134413719	-0.104301259
33	0.106735805	0.12804587	65	-0.14465155	-0.115203896
34	0.100550635	0.124811373	66	-0.154914171	-0.126261624
35	0.094499318	0.121203564	67	-0.165181495	-0.137466925
36	0.088531508	0.117238065	68	-0.175439763	-0.148808471
37	0.082567889	0.112889589	69	-0.185681902	-0.160268368
38	0.076537303	0.108171834	70	-0.195901978	-0.171827557
39	0.070402479	0.103107309	71	-0.20609542	-0.183468534
40	0.064129074	0.097699941	72	-0.216260542	-0.195173143
41	0.057677943	0.091939589	73	-0.226393699	-0.206921242
42	0.051046568	0.085825419	74	-0.236485333	-0.218692812
43	0.04427255	0.079389589	75	-0.246523325	-0.230468005
44	0.037387387	0.07268763	76	-0.256496316	-0.242227571
45	0.030394993	0.06577409	77	-0.266395583	-0.253951742
46	0.023277783	0.058677615	78	-0.276212134	-0.265621352
47	0.01602077	0.051384134	79	-0.285939277	-0.277219902
48	0.008623361	0.043889005	80	-0.295573257	-0.288731661
49	0.001072775	0.036196648			

12.5. Sample SAS Codes

The followings are the prototype of the SAS codes that will be used for the analysis. The final version of the statistical programming codes to be used will be determined prior to the database lock.

1. MMRM for the primary endpoint

```
proc mixed data = artot method = reml;
  class <usubjid> <visit> <treatment> <baseline riluzole use> <baseline edaravone use>;
  model <change in ALSFRS-R total score> = <visit> <treatment> <baseline ALSFRS-R total score>
        <baseline riluzole use> <baseline edaravone use> <visit>*<treatment>
        <visit>*<baseline ALSFRS-R total score>/ddfm=kr;
  repeated <visit> / subject = <usubjid> type = un;
  lsmeans <visit>*<treatment> / pdiff cl;
run;
```

Note: If the 2nd interim analysis result leads to an increase of the sample size, the final analyses will include the status of being included or excluded from the 2nd interim analysis as a covariate in the model.

2. Multiple imputation within each treatment group under MAR assumption

```
proc mi data=datain minimum=0 maximum=48 round=1 nimpute=50 out=dataout;
    class <baseline riluzole use> <baseline edaravone use>;
      fcs reg(<Week 4 ALSFRS-R total score> = <baseline riluzole use> <baseline edaravone use>);
      fcs reg(<Week 8 ALSFRS-R total score> = <baseline riluzole use> <baseline edaravone use>
                                              <Week 4 ALSFRS-R total score>);
      fcs reg(<Week 12 ALSFRS-R total score> = <baseline riluzole use> <baseline edaravone use>
                                               <Week 4 ALSFRS-R total score> <Week 8 ALSFRS-R total score>);
      fcs reg(<Week 16 ALSFRS-R total score> = <baseline riluzole use> <baseline edaravone use>
                                               <Week 4 ALSFRS-R total score> <Week 8 ALSFRS-R total score>
                                               <Week 12 ALSFRS-R total score>);
      fcs reg(<Week 20 ALSFRS-R total score> = <baseline riluzole use> <baseline edaravone use>
                                               <Week 4 ALSFRS-R total score> <Week 8 ALSFRS-R total score>
```

```
                                                    <Week 12 ALSFRS-R total score> <Week 16 ALSFRS-R total score>);
      fcs reg(<Week 24 ALSFRS-R total score> = <baseline riluzole use> <baseline edaravone use>
                                                    <Week 4 ALSFRS-R total score> <Week 8 ALSFRS-R total score>
                                                    <Week 12 ALSFRS-R total score> <Week 16 ALSFRS-R total score>
                                                    <Week 20 ALSFRS-R total score>);

    var <baseline riluzole use> <baseline edaravone use> <Week 4 ALSFRS-R total score> <Week 8 ALSFRS-R total score>
    <Week 12 ALSFRS-R total score> <Week 16 ALSFRS-R total score> <Week 20 ALSFRS-R total score>
    <Week 24 ALSFRS-R total score>;
run;
```

3. Multiple imputation for missing values in the reldesemtiv group with placebo distribution

   Missing data in the reldesemtiv group will be combined with all data in the placebo group for imputation, and will be filled sequentially starting at the earliest time point and progressing until the last time point. Below is an example of filling missing data at Week 8 for the reldesemtiv group using placebo data.

```
data imp; set pbo act_mis(where=(Wk8=.));
proc mi data=imp seed=&seed out=miout NIMPUTE=50;
  mcmc IMPUTE=FULL;
  var Wk4 Wk8;
run;
```

4. Tipping point analysis

```
proc mi data=datain minimum=0 maximum=48 round=1 nimpute=50 out=dataout;
    class <baseline riluzole use> <baseline edaravone use> <treatment>;
      fcs reg(<Week 4 ALSFRS-R total score> = <baseline riluzole use> <baseline edaravone use>);
      fcs reg(<Week 8 ALSFRS-R total score> = <baseline riluzole use> <baseline edaravone use>
                                              <Week 4 ALSFRS-R total score>);
      fcs reg(<Week 12 ALSFRS-R total score> = <baseline riluzole use> <baseline edaravone use>
                                              <Week 4 ALSFRS-R total score> <Week 8 ALSFRS-R total score>);
      fcs reg(<Week 16 ALSFRS-R total score> = <baseline riluzole use> <baseline edaravone use>
                                              <Week 4 ALSFRS-R total score> <Week 8 ALSFRS-R total score>
                                              <Week 12 ALSFRS-R total score>);
      fcs reg(<Week 20 ALSFRS-R total score> = <baseline riluzole use> <baseline edaravone use>
```

```
                                                    <Week 4 ALSFRS-R total score> <Week 8 ALSFRS-R total score>
                                                    <Week 12 ALSFRS-R total score> <Week 16 ALSFRS-R total score>);
        fcs reg(<Week 24 ALSFRS-R total score> = <baseline riluzole use> <baseline edaravone use>
                                                    <Week 4 ALSFRS-R total score> <Week 8 ALSFRS-R total score>
                                                    <Week 12 ALSFRS-R total score> <Week 16 ALSFRS-R total score>
                                                    <Week 20 ALSFRS-R total score>);

    var <baseline riluzole use> <baseline edaravone use> <Week 4 ALSFRS-R total score> <Week 8 ALSFRS-R total score>
    <Week 12 ALSFRS-R total score> <Week 16 ALSFRS-R total score> <Week 20 ALSFRS-R total score>
    <Week 24 ALSFRS-R total score>;
    mnar adjust (<Week 24 ALSFRS-R total score> / shift=xx adjustobs=(<treatment>='1'));
run;
```

5. Pooled estimates from the 50 imputed datasets

   For p-value from MMRM, use the combined dataset with estimates and standard errors from the 50 times of imputation

```
proc mianalyze data=est;
  modeleffects estimate;
  stderr;
run;
```

   For p-value from Wilcoxon test, apply Wilson-Hilferty transformation (Wilson & Hilferty, 1931; Goria, 1992) to the Chi-square statistic in the dataset CMH from the PROC FREQ. Note that the van Elteren test is the test comparing the rows (treatments) of the table. This is the second CMH statistic labeled "Row Mean Scores Differ", and the p-value matches the two-sided p-value from PROC NPAR1WAY (see http://support.sas.com/kb/25/022.html).

```
proc npar1way wilcoxon;
  strata <stratum1> <stratum2>;
  class <treatment>;
  var <rank>;
run;

proc freq;
  table <stratum1>*<stratum2>*<treatment>*<rank> / cmh scores=modridit noprint;
  ods output cmh=cmh;
run;
```

```
Data cmh_combine; set cmh;
If AltHypothesis='Row Mean Scores Differ';
chi_value_wh=((Value/DF)**(1/3) - (1-2/(9*DF)))/SQRT(2/(9*DF));
chi_sterr_wh=1;
run;
```

Use the combined dataset with the transformed statistics from the 50 times of imputation

```
Proc mianalyze data=cmh_combine;
 modeleffects chi_value_wh;
 stderr chi_sterr_wh;
run;
```

Signature Manifest

Document Number: PRD-0314 Revision: 00

Title: CY 5031 - Statistical Analysis Plan - Version 2.0 - A Phase 3, Multi-Center, Double-Blind, Randomized, Placebo-Controlled Trial to Evaluate the Efficacy and Safety of Reldesemtiv in Patients with Amyotrophic Lateral Sclerosis (ALS)

Effective Date: 27 Feb 2023

All dates and times are in Pacific Time.

CY 5031 - Statistical Analysis Plan - Version 2.0 - A Phase 3, Multi-Center, Double-Blind, Randomized, Placebo-Controlled Trial to Evaluate the Efficacy and Safety of Reldesemtiv in Patients with Amyotrophic Lateral Sclerosis (ALS)

1: Electronic Approvals

Name/Signature	Title	Date	Meaning/Reason
Jenny Hsiao-Lan Wei (JWEI)	Director, Biostatistics	27 Feb 2023, 01:39:08 PM	Approved
Stacy Rudnicki (SRUDNICKI)	Vice President, Clinical Research	27 Feb 2023, 01:50:28 PM	Approved
Stuart Kupfer (SKUPFER)	SVP, Chief Medical Officer	27 Feb 2023, 02:45:25 PM	Approved
Lisa (Lixin) Meng (LMENG)	Vice President, Biometrics	27 Feb 2023, 04:10:30 PM	Approved