Confidential

Japanese Early-stage Trial of high-dose methylcobalamin for Amyotrophic Lateral Sclerosis (JETALS)

Investigator-Initiated Clinical Trials
(Protocol Number: E0302-TOK-763)

Statistical Analysis Plan (Draft)
Statistical analysis plan
Version 2.1

Date: 24 March 2020

Supervisor for statistical analysis: Tatsuo Kagimura

Kobe Medical Industry and Urban Promotion Organization
Medical Innovation Promotion Center

Translational Research Center for Medical Innovation
Founded in 2003 by MEXT & Kobe City
For acceleration of translational research in Japan

Contents

106 List of abbreviations

Abbreviations	Non-abbreviated expressions (English)
ADS	Analysis dataset
ALS	Amyotrophic lateral sclerosis
ALSAQ-40	ALS assessment questionnaire-40
ALSFRS-R	ALS Functional Rating Scale-Revised
ATC classification	Anatomical Therapeutic Chemical Classification System
BMI	Body mass index
CI	Confidence interval
FAS	Full analysis set
FVC	Forced vital capacity
GCP	Good clinical practice
LLT	Lowest level term
LSMean	Least square mean
MedDRA/J	Medical Dictionary of Regulatory Activities/J
MMRM	Mixed effect model for Repeated Measures
MMT	Manual Muscle Testing
MRC score	Medical Research Council score
PMDA	Pharmaceuticals and Medical Devices Agency
PPS	Per Protocol Set
PT	Preferred Term
Q1	25 percentile
Q3	75 percentile
SD	Standard deviation
SE	Standard error
SOC	System Organ Class
TRI	Translational Research Informatics Center for Medical Innovation
WHO DD	WHO Drug Dictionary

107
108

1 BASIC ANALYSIS-RELATED CONTENT

1.1 Objectives and Research Question of the Study

Objective:
To evaluate the superiority of high-dose intramuscular E0302 (mecobalamin 50 mg) over placebo in patients with amyotrophic lateral sclerosis (ALS) using the Japanese-language revised ALS functional rating scale (ALSFRS-R) scores as indicators. We will also evaluate the safety of high-dose intramuscular E0302 (Methylcobalamin).

Research Question:
Verification proposition: Can high-dose intramuscular mecobalamin (50 mg) when compared with placebo prevent the progression of symptoms at 16 weeks using the ALSFRS-R as an index?

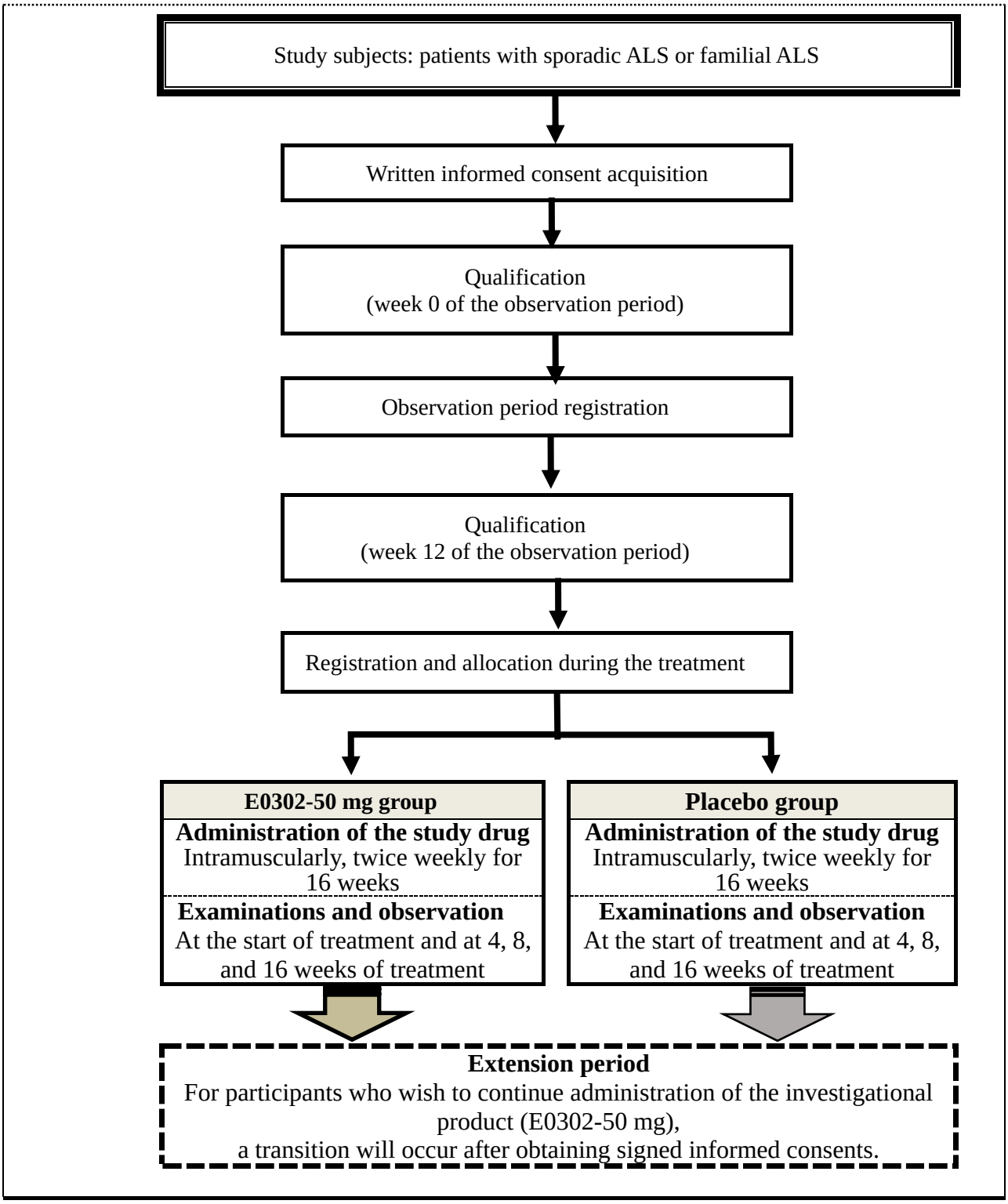
Estimand;

1. Target population:
2. Variables (or Evaluation items): Maximum analytic population (defined by analytic population).
3. Consideration of intermediate events: Change in ALSFRS-R total score between baseline to week 16 (defined by primary outcome).
   Measurement after cessation of treatment: Measurements taken after the date of discontinuation are treated as missing data.
   Measurement after administration of edaravone (prohibited drugs from concomitant use): treated as missing data.
   Measurement after administration of new dose or increased dose of riluzole (limited drugs from concomitant use): treated as a missing data.
4. Summary of variables at the population level: Reject the null hypothesis that the means of the differences between groups are equal.

Research Strategy Positioning:
Previous research: Phase II/III study (E0302-J081-761)
Three-arm comparative study of placebo and E0302 at 25 and 50 mg. In the subpopulation of patients within 1 year of onset with a decrease of 1 to 2 points in the ALSFRS-R during the observation period (12 weeks) in this study, the mean ± standard deviation of the change in total ALSFRS-R score at 16 weeks was -3.2 ± 4.0 points in the mecobalamin 50 mg group (n=26) and -5.8 ± 5.0 points in the placebo group (n=32) (mean difference: -2.6).

144

145 **1.2 Schema**



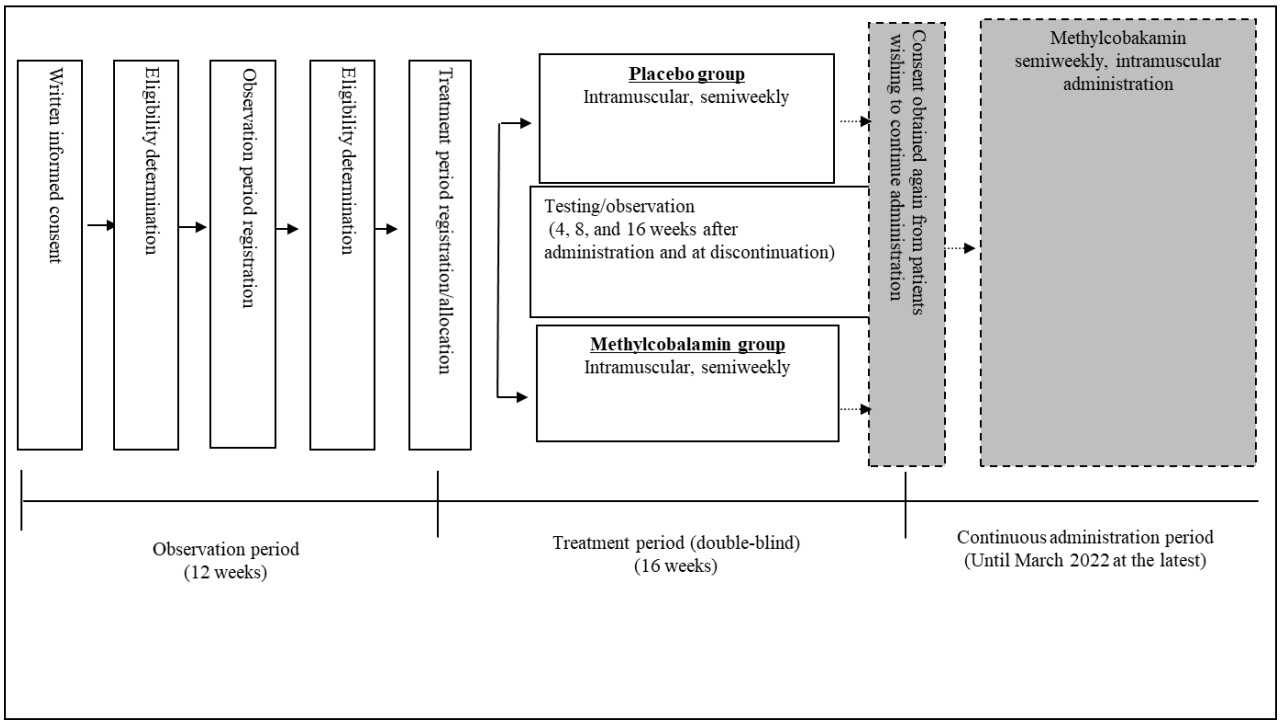
146
147

1.3 Summary of studies

Target sample size	128 patients (64 in the placebo group, 64 in the E0302-50 mg group)
Case registration period	2017.10-2019.9 (2 year)
Date of completion of follow-up	2020.3
Design	Double-blind, parallel-group study
Allocation	The following items will be used as allocation adjustment factors (minimization factors) for allocation using a variant of the minimization method, taking into account the balance within the institution and the overall balance of the treatment groups: ・Disease type (bulbar, upper extremity, and lower extremity) ・Severity of ALS at the end of the observation period (grade 1 and 2) ・Time from onset to start of the observation period (≤9 months, 9 to 12 months) ・%FVC at the end of the observation period (<90%, or ≥90%) ・Previous administration of edaravone (yes/no)

Observation flow diagram:

Trial design



154 Study schedule

Timing		Observation period		Treatment period (double-blinded)			Contunious treatment period		Disontinuation (Treatment period/ Continuous treatment period)
		Week 0 Initiation	12 weeks (end)	Week 0 (Initial day of administration*1)	4 weeks	8 weeks	16 weeks (Completed/Discontinued)	Every 12 weeks (Completed/Discontinued)	
Acceptable range (weeks)		—	±1				—	±2	±2*14
				—	±1	±1	±2		
Tests, observations, and assessments	Obtainment of written informed consent	●					●*2		
	Patient Characteristics	●	●						
	Registration of the observation period	●							
	Enrollment and assignment during the Treatment period		●						
	Diagnosis (including electromyography and nerve conduction studies)	●*11	●						
	Event occurrence*3			←	―	―	―	―	→
	ALSFRS-R*12	●	●		●	●	●	●	●
	%FVC	●	●			●	●		
	MMTs; handgrip strength testing; Norris scale; ALSAQ-40		●*4			●	●		
	Blood homocysteine (intensive		●				●		
	Clinical Laboratory Tests (Hematology, Biochemistry, Urine)	●*5	●*6		●*7		●*7	●	●
	12-lead ECG	●	●*8	●* 8, 9			●*9, *10	●	
	Vital signs	●	●		●	●	●	●	●
	Administration of the investigational drug			With E0302 - 50 mg/dose or placebo Intramuscular injection twice weekly*13			E0302 - 50 mg/dose Intramuscular injection twice weekly*13		
	Treatment conditions for investigational drug			←	―	―	―	―	→
	Tracheostomy status			←	―	―	―	―	→
	Concomitant medication/treatment *12	←	―	―	―	―	―	―	→
	Adverse events*3	←	―	―	―	―	―	―	→

*1 Should be conducted from Day 0 (Allocation date) to Day3. *2 Obtained informed consent from Week 12 to Week 16. *3 Occurrence of the Event and Adverse event are required to be reported from the last administration date to 28 days later. *4 Conducted in the eligible patients for the treatment period. *5 Measurement of serum Vitamin B12 level is conducted. *6 Women only take a pregnancy test. *7 Conducted before administration of the investigational product. *8 QT assessments are conducted before administration of the investigational product if the first administration is on the allocation date. *9 QT assessments are conducted twice (before administration and 2 hours later after administration). *10 Conducted from Week 8 to the last administration date of Week 15. *11 The results of electromyography and nerve conduction studies conducted in the other medical institutions allowed to be evaluated. *12 Among the patients enrolled in the treatment phase, ALSFRS-R evaluation and investigation of concomitant medication and concomitant therapy will be conducted as much as possible until Week 16 of the treatment phase for the discontinued patients except for the untreated patients (except for those who refuse to continue participation in the study or withdraw consent). *13 For twice-weekly administration of the study drug, the drug will be administered twice during a 7-day period starting from the start date of the treatment/continuation period. The interval between doses should be at least one day, and two doses (4 vials) on the same day is not allowed. ※14 Performed after last administration of the investigational drug and within ±2 weeks from the discontinuation date.

155

1.4 Data storage location

The data provided by the DM and the analysis data set (ADS) used for the analysis will be saved as follows:

DM Dataset:	Z:¥01_プロジェクト¥TRI プロジェクト 2¥325_TRINEU1701(梶)¥09_生物統計¥09_最終解析¥2_ANALYSIS¥EXECUTE1¥01_DATA¥DM_RAW
ADS:	Z:¥01_プロジェクト¥TRI プロジェクト 2¥325_TRINEU1701(梶)¥09_生物統計¥09_最終解析¥2_ANALYSIS¥EXECUTE1¥01_DATA¥ADS

1.5 Analysis software

Analyses will be performed using software packages SAS Version 9.4 (SAS Institute, Cary, NC, USA) and R version 3.6.

2 Analysis purpose

This statistical analysis plan defines the details of the final analysis specified in the protocol.

2.1 Changes in Analytical Procedures from the Protocol

Changes in analytical methods from Protocol version 5.0 (dated April 24, 2019) are shown below.

Descriptions after change	Reasons for change
18.2.2.1 Analysis of subject characteristics Add unpaired t-test in addition to Wilcoxon rank-sum test and Fisher's exact test for comparison of homogeneity between treatment groups.	In the previous trial (E0302-J081-761), an unpaired t-test was adopted. The same test should be used to compare the homogeneity for similar characteristic items.
18.2.2.1 Analysis of subject characteristics Summary statistics (mean, standard deviation, minimum, median, and maximum) of blood homocysteine concentration at week 12 of the observation period will be estimated and compared by the corresponding t-test between eligible and ineligible patients for the treatment period.	It is necessary to confirm whether blood homocysteine concentration is similar or not between eligible and ineligible patients for the treatment period.
18.2.2.3 Prior and Concomitant Therapies There will be no data collection by PPS.	PPS is not the primary analysis, and there is no point in comparing the two groups.
18.2.2.4.1 Primary Endpoint A secondary analysis comparing the change at each time point between groups by the Wilcoxon rank-sum test will be added.	For comparison with the results of the previous trial (E0302-J081-761), the same test method as in the last phase will be used for intergroup comparison.
18.2.2.4.1 Primary Endpoint The following analyses will be added as sensitivity analyses. The ALSFRS-R total score at baseline, week 4, 8, and 16 will be used as the response variable, and a linear regression equation will be fitted to the time point and response variable, and the slope between groups will be compared by a mixed model with the intercept and slope as the variable effects.	To estimate the pattern of ALSFRS-R total scores over time and to compare between groups.
18.2.2.4.1 Primary Endpoint The following analysis will be added as a supplementary analysis. (1) Primary and secondary analysis will be performed. Missing values due to death were assigned the previous value and the worst possible value (0 point). (2) The AUC of the change in ALSFRS-R from baseline to week 16 will be used as the	(1) The primary analysis will not compensate for missing data due to death during the treatment period. Therefore, we will conduct a supplementary analysis under some assumptions about missing data. (2) To examine the average transition from the baseline to week 16. (3) It is necessary to conduct an analysis with different handling of intermediate events to

endpoint, and the groups will be compared by analysis of covariance with the baseline value as the covariate and the treatment group as the factor. (3) Primary and secondary analyses will be performed, which include the data from patients who are administered edaravone or changed the daily dose of riluzole during the treatment period.	confirm the robustness of the analysis results.
18.2.2.4.1 Primary Endpoint Summary statistics will be calculated for each stratum of the stratification factors (4.4.1), and the change at each time point will be compared between groups by the Wilcoxon rank-sum test.	To perform stratified analysis to explore differences in efficacy across patient populations.
18.2.2.4.1 Primary Endpoint ALSFRS-R sub-score (Bulbar score, Limb score, and Respiratory score) will be compared between groups by calculating summary statistics and Wilcoxon rank-sum test for change at each time point.	To explore what factors are responsible for the changes in ALSFRS-R, we will examine the changes in the ALSFRS-R sub-score.
18.2.2.4.2 Secondary Endpoint Norris scale sub-score Limb Norris Scale score Norris Bulbar Scale score ALSAQ-40 sub-score Physical Mobility ADL Eating and Drinking Communication Emotional Functioning	To explore what factors are responsible for the changes in the Norris scale and ALSAQ-40, we will examine the changes in their sub-score.
18.2.3.1.Adjustment Using Covariates The following analyses will not be performed. Multiple regression analysis will be performed to determine the impact on the primary endpoint of background and prognostic factors that were not found to be homogenous between treatment groups among the subject characteristics that may affect the clinical evaluation of ALS (to be determined prior to key opening).	Multiple regression analysis to determine the impact on the primary analysis of the background and prognostic factors that were not found to be homogenous between treatment groups were replaced with the sensitivity analysis of the primary analysis. This analysis directly reinforces the testing hypothesis of the study.

3 Analysis Sets

Of the patients enrolled during the treatment phase, we will include individuals meeting the inclusion criteria in the safety analysis population, and those who do not meet the inclusion criteria will be excluded:

- Individuals from whom we have not obtained signed informed consent forms.
- Individuals who do not meet the inclusion criteria for the primary target disease (inclusion criteria [1] to [5]).
- Individuals who have not received the investigational product.
- Individuals with no evaluable safety data.

3.1 Full Analysis Set

We will exclude the following individuals from the full analysis set (FAS) during the enrollment in the treatment phase:

- Individuals who do not meet the primary inclusion criteria (inclusion criteria [1] to [5]).
- Individuals who are categorized under a GCP violation, such as administration outside the study period or not providing signed informed consent forms.
- Individuals lack data on the components of the primary endpoint (ALSFRS-R).
- Individuals who have not received the investigational product.

3.2 Protocol-compliant population (PPS)

The Per Protocol Population (or per-protocol set, PPS) is defined as the set excluding individuals who meet the following criteria from the FAS:

- Individuals lack data from week 8 onward regarding the component of the primary endpoint that can be assessed (ALSFRS-R).
- Individuals who meet exclusion criteria affecting efficacy assessment (exclusion criteria [1], [8], [14]).
- Individuals who have a cumulative injection rate lower than 70% of the investigational product until the date of completion or discontinuation of the study.
- Individuals who withdraw from the study after less than 8 weeks.

3.3 Significant protocol deviations

Major protocol deviations include the following:

- Individuals lack data from week 8 onward regarding the component of the primary endpoint that can be assessed (ALSFRS-R).
- Individuals meeting exclusion criteria affecting efficacy assessment (exclusion criteria [1], [8], [14]).
- Individuals who have a cumulative injection rate lower than 70% of the investigational product until the date of completion or discontinuation of the study.
- Individuals who withdraw from the study within 8 weeks.

Researchers at the data center will evaluate protocol deviations for case accrual and define critical protocol deviations and case management during case review meetings by the time the study data are fixed.

4 Definition of endpoints

The endpoints of the double-blind study period will be as follows.
The endpoints of the continuous treatment period will be the time from baseline to the date of an event (all-day use of non-invasive respiratory support, use of invasive respiratory support, or death), the change in ALSFRS-R total score, adverse events, laboratory test values, vital signs, and electrocardiograms.

4.1 Efficacy endpoint

4.1.1 Primary endpoint

- Change in ALSFRS-R total scores from baseline to 16 weeks of treatment.

The change from baseline in ALSFRS-R total scores at weeks 4 and 8 of the treatment period will be the secondary analysis endpoint.
ALSFRS-R total scores at different time points during the observation period: week 0, 12 of the observation period, week 4, 8, and 16 (completion/discontinuation) of the treatment period).

The following ALSFRS-R itemized scores measured at each time point will also be analyzed as supplement data.

- Bulbar score: Total scores of Speech, Salivation, and Swallowing.
- Limb score: Total scores of Handwriting, Cutting food, Dressing, and hygiene, Turning in bed, Walking, and Climbing stairs.
- Respiratory score: Total scores of Dyspnea, Orthopnea, and Respiratory insufficiency.

The area under the curve (AUC) for the double-blind study period of the change in ALSFRS-R total score is calculated by the following formula. However, if the patient does not complete the 16-week treatment period, the weekly AUC up to the last measurement point should be calculated.

$$AUC_{DBT} = \frac{(0 + R_4) \times 4 + (R_4 + R_8) \times 4 + (R_8 + R_{16}) \times 8}{2 \times 16}$$

R_4: Change in ALSFRS − R total score at week 4 from the last score
R_8: Change in ALSFRS − R total score at week 8 from the last score
R_{16}: Change in ALSFR − R total score at week 16 from the last score

4.1.2 Secondary endpoint

4.1.2.1 Time-to-event

Time from the investigational product allocation date to the first occurrence of any of the following events:

- Full-day non-invasive respiratory support
- Use of invasive respiratory support device
- Death

The time-to-event is determined by the following equation:
= investigational drug administration start-day + 1

4.1.2.2 Change in %FVC

Measurements at each time point during week 0 (start) and week 12 (end, baseline) of the observation period, week 8, and week 16 (completion/discontinuation) of the treatment period. Calculate the change from baseline to week 8 and week 16 (at discontinuation) of the treatment period at each time point.

4.1.2.3 Change in blood homocysteine levels

Intensive measurements at week 12 (end) and week 16 (completion/discontinuation) in the observation period. Calculate the change from baseline to week 16 (at discontinuation) of the

treatment period at each time point.

4.1.2.4 Change in MMT total score

Measurements at each time point during week 12 (end, baseline) of the observation period, week 8, and week 16 (completion/discontinuation) of the treatment period. Calculate the change from baseline to week 8 and week 16 (at discontinuation) of the treatment period at each time point.

4.1.2.5 Change in Grip strength (right and left)

Measurements at each time point during week 12 (end, baseline) of the observation period, week 8, and week 16 (completion/discontinuation) of the treatment period. Calculate the change from baseline to week 8 and week 16 (at discontinuation) of the treatment period at each time point.

4.1.2.6 Change in Norris scale total score

Measurements at each time point during week 12 (end, baseline) of the observation period, week 8, and week 16 (completion/discontinuation) of the treatment period. Calculate the change from baseline to week 8 and week 16 (at discontinuation) of the treatment period at each time point.
Measurements of the following Norris scale sub-scores at each time point will also be analyzed as a supplement.

- Limb Norris Scale total score
- Norris Bulbar Scale total score

4.1.2.7 Change in ALSAQ 40 total score

Measurements at each time point during week 12 (end, baseline) of the observation period, week 8, and week 16 (completion/discontinuation) of the treatment period. Calculate the change from baseline to week 8 and week 16 (at discontinuation) of the treatment period at each time point.
Measurements of the following ALSAQ-40 sub-items at each time point will also be analyzed as a supplement.

- Physical Mobility [10 items: 1-10]
- ADL (Activities of Daily Living), Independence [10items: 11-20］
- Eating and Drinking [3 items: 21-23]
- Communication [7 items: 24-30]
- Emotional Functioning [10 items: 31-40]

4.2 Other efficacy endpoints

None

4.3 Safety endpoint

4.3.1 Adverse events

Adverse events reported will be assigned a Lowest Level Term (LLT) code using the MedDRA/J dictionary. The MedDRA and J versions used for analysis will be updated at the time of database locking.
Adverse events will be summarized for the period from post-study drug administration to week 16 of treatment (completion/discontinuation). Adverse events in the double-blind study period for patients who have progressed to the continuous-dose phase are those that occur up to the day before the first dose of the continuous treatment period.
Adverse reactions will be considered “related” to the study drug.

The starting date for the time-to-onset of the adverse event is the starting day of administration.
Time-to-onset of adverse event = [date of onset of adverse event] - [study drug start date] + 1

If more than one adverse event of the same preferred term (PT) occurs in a single subject, we will count the number of adverse events as one subject. If a causal relationship is judged to be causal

even once, the causal judgments of the cases will be aggregated as causal.

4.3.2 Laboratory test values

Measurements at each time point during week 0 (start) and week 12 (end, baseline) of the observation period, week 4 and week 16 (completion/discontinuation) of the treatment period, and the continuous treatment period.

4.3.3 Vital signs

Measurements at each time point during week 0 (start) and week 12 (end, baseline) of the observation period, week 4 and week 16 (completion/discontinuation) of the treatment period, and the continuous treatment period.

4.3.4 Electrocardiogram

Measurements for heart rate, RR interval, PR interval, RQS width, QT interval, QTcB, QTcF, central reading QTcB, and central reading QTcF twice before and after treatment at each time point during week 0 and week 16 (completion/discontinuation) of the treatment period.
ECG abnormalities at week 0 (start) and week 12 (end, baseline) of the observation period, and week 0 and week 16 (completion/discontinuation) of the treatment period.

4.4 Other items

4.4.1 Stratification factor

- Age at onset: <65 years, ≥65 years
- Sex: male, female
- Disease type: aggregate in the following two categories
  1) bulbar, limb (upper extremity or lower extremity).
  2) Bulbar, upper extremity, lower extremity.
- Time from onset to start of follow-up: 9 months or less, greater than 9 months, and 12 months or less.
- %FVC at the end of observation period: less than 90%, more than 90%.
- Change in %FVC during the observation period: <0, ≥0
- History of edaravone administration: yes, no
- Concomitant use of riluzole: yes, no
- Age at onset: aggregate in the following four categories
  1) <65 years, ≥65 years
  2) <60 years, ≥60 years
  3) <70 years, ≥70 years
  4) <50 years, 50-59 years, 60-69 years, ≥70 years
- Nutrition status: BMI <18.5, ≥18.5
- The updated Awaji criteria at the end of the observation period: Definite, Probable, Probable laboratory-supported
- The updated Awaji criteria at the end of the observation period: Definite, Probable, Probable laboratory-supported
- The El Escorial revised Airlie House diagnostic criteria at the end of the observation period: Definite, Probable, Probable laboratory-supported
- Muscle strength of neck flexors at the end of observation period: MRC score 5, ≤4
- ALS severity grade at the end of observation period: Grade1, 2
- Change in ALSFRS-R total score during the observation period: -1, -2 points
- ALSFRS-R total score at the end of observation period: ≤37, 38‐42, ≥43

A subpopulation of ALS patients who are diagnosed with Definite, Probable, Probable-laboratory supported grade by the El Escorial revised Airlie House diagnostic criteria.

- The diagnostic criteria of ALS in this study (the updated Awaji criteria) have been modified from the El Escorial revised Airlie House diagnostic criteria in the previous study (E0302-J081-761). For comparison with the previous study, we will conduct an analysis in a subpopulation using the same diagnostic criteria as in the previous study.

Additional categories for safety analysis:

- Grade by the degree of liver dysfunction at screening: Normal, Grade 1, ≥2
- Grade by the degree of renal dysfunction at screening: Normal, Grade 1, ≥2

4.4.2 Cumulative administration rate

- Cumulative administration rate = [Number of administration (times)] / [Prescribed administration period (weeks) *2] *100

Prescribed administration period (weeks) = ([Date of last administration] - [Initiation date of administration] + 1) / 7
The decimal point of the prescribed administration period is rounded up. The prescribed administration period for the double-blind period does not exceed 16. For patients who experienced an event, the date of administration prior to the date of the event is used as the last date of administration.

5 Data handling

5.1 Baseline

The baseline is the observation period of week 12 (at the end of the observation period).

5.2 Time window

The time point of the observation period in the analysis of the efficacy endpoints will be the calculated visit, in which the time point of the visit is calculated from the end of the treatment period by the following time window.
Number of observation days are as follows: [Number of observation days] = [Date of evaluation] - [Initiation date of administration] + 1
If two observed values are found within the same time window, the one closer to the target number of days is used as the observed value at that time.

Visit	Target weeks	Allowance of protocol	Allowance of analysis	The target number of days	The lower limit of days	The upper limit of days
Baseline	12	1	1	84	77	91
Week 4 (the treatment period)	4	1	1	28	21	35
Week 8 (the treatment period)	8	1	2	56	42	70
Week 16 (the treatment period)	16	2	2	112	98	126
Week 12 (the continuous administration period)	28	2	4	196	168	224
Week 24 (the continuous administration period)	40	2	4	280	252	308
Week 36 (the continuous administration period)	52	2	4	364	336	392
Week 48 (the continuous administration period)	64	2	4	448	420	476
Week 60 (the continuous administration period)	76	2	4	532	504	560
Week 72 (the continuous administration period)	88	2	4	616	588	644
Week 84 (the continuous administration period)	100	2	4	700	672	728
Week 96 (the continuous administration period)	112	2	4	784	756	812
Week 108 (the continuous administration period)	124	2	4	868	840	896
Week 120 (the continuous administration period)	136	2	4	952	924	980

5.3 Handling of measured values at the administration of prohibited drugs and restricted concomitant drugs

- The measured values after administration of edaravone (prohibited drug) are handled as missing values.
- The measured values after administration of increased dose or a new start of riluzole (restricted concomitant drug) are handled as missing values.

5.4 Completion of missing values

For all assessment items, missing data values will not be imputed.

5.5 Weekly and monthly annual calculations

We will calculate weeks and months from days according to the following formulas:
Weekly conversion: days/7
Monthly conversion: days × (12/365.25)
Annual conversion: days/365.25

5.6 Handling of laboratory value detection limits

In our summary statistics of laboratory values, we will handle the laboratory values (the detection limit of laboratory values) as the laboratory limit measurements. However, a list with missing data will indicate it is below the quantitation limit.

6 Analysis method

The FAS analysis will be the primary efficacy analysis, and the PPS analysis will be performed as an additional sensitivity analysis for the primary and secondary endpoints.

- In the calculation of summary statistics, the decimal point of the mean, standard deviation (SD), and median will be expressed as an increase of 1 digit. Similarly, the decimal points of adjusted means (LSMean) and SEs will also be presented as increasing by orders of magnitude. The decimal point of the minimum and maximum values will be the same.
- The level of significance in a test will be 5% two-sided.
- For p values, a value lower than 0.001 will be labeled "p <0.001". In the case of 0.001 or more, the four digits after the decimal point will be rounded off, and the indicated digits will be three digits after the decimal point.
- All calculations of confidence intervals will be two-sided 95% confidence intervals.
- The number of cases will be expressed as integers, and the percentages of cases and their 95% confidence intervals will be shown to one decimal place. We will calculate the 95% confidence interval for the incidence (%) using the Clopper-Pearson method with accurate confidence intervals.

6.1 Subject disposition

We will prepare flow diagrams for patients enrolled during the observation period, discontinued during the observation period, enrolled during the treatment period, untreated, discontinued during the treatment period, and for those who complete the treatment period. Cases of discontinuation in the observation period, untreated cases, and discontinuation in the treatment period will be tabulated by reason in a list.

6.2 Sample for analysis

We will present the number of patients for each SAS, FAS, and PPS. The exclusions will be tabulated by reason for exclusion in an exclusion list.

6.3 Patient characteristics

For patient demographics, we will determine summary statistics (sample size, mean, SD, minimum, median, and maximum) for quantitative variables and sample size/percentage for each treatment group for qualitative variables. We will determine bias between groups by Fisher exact test, unpaired t-test, and Wilcoxon rank-sum test. Categories of background items will be shown in the sample figures and tables (SAS, FAS, PPS).
The frequency of complications and medical history will be summarized by system organ class (SOC: System organ class) and PT using the MedDRA/J dictionary (SAS).
We will assign concomitant medications WHO Drug Global codes and will aggregate frequencies by the group for each ATC classification during the treatment period and the continuous treatment period (SAS, FAS).
The frequency of combination therapy will be summarized by the group for each combination treatment during the treatment period and the continuous treatment period (SAS, FAS).
Summary statistics (mean, SD, minimum, median, and maximum) will be compared between the treatment period enrollees and treatment period ineligible individuals by paired *t*-test for homocysteine measurements at week 12 of the observation period.

6.4 Treatment status

Summarize the number of administrations, duration of treatment, administration rate (number of

administrations/scheduled administrations), number of drug withdrawal, days of drug withdrawal, and treatment discontinuation rate for each group during the double-blind period (SAS, FAS, PPS). Summarize the number of administrations, duration of treatment, treatment discontinuation rate, and rate of self-administration (numbers of self-administration / numbers of administrations) during the continuous treatment period (SAS, FAS).

6.5 Efficacy endpoint

6.5.1 Primary endpoint

For FAS, the change from baseline in the total ALSFRS-R score at 4, 8, and 16 weeks will be fitted to a mixed-effects model with baseline values as covariates, treatment group, time points at 4, 8, and 16 weeks, minimizers (disease type, severity of ALS at the end of the observation period, time from initial onset to start of the observation period, %FVC at the end of the observation period, and history of edaravone treatment), the interaction between time point and treatment group as a population effect, and patient as a variable effect. We will fit a mixed-effects model with repeated measurements (MMRM) that include the interaction between time point and treatment group as a population effect and the patient as a variable effect.

For the estimation of the variance-covariance matrix between time points, we assume an unstructured structure based on the data, but if it does not converge, we use Heterogeneous TOEP (TOEPH), Heterogeneous AR(1) (ARH(1)), Heterogeneous CS (CSH) Toeplitz (TOEP), Spatial Power (SP(POW)), Autoregressive (1) (AR(1)), and Compound Symmetry (CS).

The missing values at each time point will not be compensated. The time point of administration, including treatment discontinuation cases, will be calculated from the observation date of ALSFRS-R in the time window shown in Section 5.2.

Primary Analysis:
According to the model described above, the upper limit of the 95% confidence interval of the least-squares means (LSMean) of the difference between the placebo group and the E0302 50 mg group (placebo group - E0302 50 mg group) of the change in the total number of points in the ALSFRS-R at week 16 of the treatment period will be considered significant if it is below 0.

Secondary analysis:

- The LSMean and its 95% confidence interval for each group at 4, 8, and 16 weeks of the treatment period and the LSMean and its 95% confidence interval for the difference between groups and the P-value for the test of difference (not adjusted for multiplicity) will be calculated.
- A longitudinal trend graph for each group with the mean at each time point on the vertical axis and the time point on the horizontal axis will be created.
- Summary statistics (number of patients, mean, standard deviation, minimum, first quartile, median, third quartile, and maximum) of the measurements and changes at each time point from the beginning of the treatment period to 16 weeks will be calculated.
- Intergroup comparisons will be made by the Wilcoxon rank-sum test for change at each time point (for comparison with study E0302-J081-761).

Sensitivity analysis:

- Among the subject characteristics that are considered to affect the clinical evaluation of ALS, the target disease of this study (Section 4.4.1 Stratification Factors), the analysis will be conducted using an analytical model with background items for which uniformity between treatment groups could not be confirmed and prognostic factors considered necessary added as adjustment factors.
- If the primary analysis does not converge on an unstructured structure, an analysis will be conducted in which the minimizing factors are excluded from the analysis model.
- The ALSFRS-R total score at 0 (baseline), 4, 8, and 16 weeks of the treatment period will be used as the response variable, and a first-order regression equation will be fitted to the time point and response variable, and the slope between groups will be compared using a mixed model with the intercept and slope as variable effects.

Supplemental analysis:

- A similar primary and secondary analysis and sensitivity analysis will be performed for PPS.
- A similar primary and secondary analysis will be performed with missing values due to death compensated by the previous value.
- A similar primary and secondary analysis will be performed with missing values due to death compensated for by the worst value of zero.
- Intergroup comparisons will be performed by analysis of covariance with the baseline value as the covariate and the treatment group as the factor, using the AUC of change through week 16 as the endpoint.
- The same primary and secondary analyses will be performed for patients who started taking edaravone or riluzole during the treatment period or who changed the daily dose of riluzole, including measurements after the new or increased dose.
- The same primary and secondary analysis will be performed for the subpopulation of clinically definite ALS, clinically probable ALS, and clinically probable-laboratory-supported ALS according to El Escorial revised Airlie House diagnostic criteria at the end of the observation period.
- Summary statistics will be calculated for each stratum of the stratification factors (Section 4.4.1), and the groups will be compared by the Wilcoxon rank-sum test for change at each time point.
- Summary statistics will be calculated for each stratum of the ALSFRS-R for spherical function, motor function of extremities, and respiratory function, and the change at each time point will be compared between groups using the Wilcoxon rank-sum test.

Analysis of the continuous treatment period data:

The following analyses will be performed on the ALSFRS-R from the beginning of the treatment period to the end of the continuation period (Cut off data in case of continued administration).

- Summary statistics (number of patients, mean, standard deviation, minimum, first quartile, median, third quartile, and maximum) of the measurements and changes from baseline to each time point will be calculated for the whole and each group. The amount of change will be tested by the signed-rank test. The same analysis will be performed for the item-specific scores of the ALSFRS-R for bulbar function, motor function of limbs, and respiratory function.
- The median value at each time point for each group and the entire group will be used as the vertical axis, and the time point will be used as the horizontal axis to create a trend graph over time.

6.5.2 Secondary endpoint

The following analyses will be performed for FAS and PPS.

6.5.2.1 Time-to-event

The Kaplan-Meier method will be used to generate survival curves for each group and for the entire study, and the cumulative survival rate and its 95% confidence interval will be calculated for 16 weeks of the treatment period and each 12-week period of the continuous treatment period. The hazard ratios and their confidence intervals will be calculated by the proportional hazards model, using the log-rank test for the time to event up to 16 weeks, with censoring after 16 weeks.

6.5.2.2 %FVC changes

Analyses similar to the primary and secondary analyses of the ALSFRS-R total score (Section 6.5.1) will be performed. A supplementary analysis will be performed in PPS and the dataset in which missing values due to death were assigned the previous value.

6.5.2.3 Blood homocysteine level changes

Analyses similar to the primary and secondary analyses of the ALSFRS-R total score (Section 6.5.1) will be performed. A supplementary analysis will be performed in PPS and the dataset in which missing values due to death were assigned the previous value.

6.5.2.4 MMT total score changes

Analyses similar to the primary and secondary analyses of the ALSFRS-R total score (Section 6.5.1) will be performed. A supplementary analysis will be performed in PPS and the dataset in which missing values due to death were assigned the previous value.

6.5.2.5 Grip strength changes (right and left)

Analyses similar to the primary and secondary analyses of the ALSFRS-R total score (Section 6.5.1) will be performed. A supplementary analysis will be performed in PPS and the dataset in which missing values due to death were assigned the previous value.

6.5.2.6 Sum of Norris scale changes

Analyses similar to the primary and secondary analyses of the ALSFRS-R total score (Section 6.5.1) will be performed. A supplementary analysis will be performed in PPS and the dataset in which missing values due to death were assigned the previous value. Summary statistics will be calculated for the sub-items, and the groups will be compared by the Wilcoxon rank-sum test for the amount of change at each time point.

6.5.2.7 ALSAQ 40 total score changes

Analyses similar to the primary and secondary analyses of the ALSFRS-R total score (Section 6.5.1) will be performed. A supplementary analysis will be performed in PPS and the dataset in which missing values due to death were assigned the previous value. Summary statistics will be calculated for the sub-items, and the groups will be compared by the Wilcoxon rank-sum test for the amount of change at each time point.

6.5.3 Other Efficacy Endpoints

None

6.6 Safety evaluation

6.6.1 Adverse event

The following analysis of adverse events in the double-blind period will be performed.

- The number of cases and percentage of incidence of each group will be tabulated to summarize the events related to adverse events. The 95% confidence intervals for the incidence of adverse

events, adverse reactions, serious adverse events, and adverse events leading to discontinuation will be calculated and compared between the placebo group and the E0302 50 mg group using Fisher's direct probability test.

- The number and percentage of adverse events will be tabulated for each group, SOC, and PT.
- The number and percentage of side effects will be tabulated for each group, SOC, and PT.
- The number and percentage of adverse events will be tabulated for each group, SOC, PT, and severity.
- The number and percentage of side effects will be tabulated for each group, SOC, PT, and severity.
- The number and percentage of severe adverse events will be tabulated for each group, SOC, and PT.
- The number and percentage of incidents of adverse events will be tabulated for each group, serious/nonserious, SOC, and PT.
- The number and percentage of incidents of adverse events leading to discontinuation of the investigational drug will be tabulated for each group, SOC, PT, and causal relationship.
- The incidence of adverse events and the incidence of adverse drug reactions in the subpopulation of the stratification factors (Section 4.4.1) will be tabulated for each group. The initial symptoms will be tabulated by two categories: bulbar onset and limb onset, and the age of onset will be tabulated by two categories: under 60 years old and over 60 years old.
- We will prepare a list of all deaths, a list of serious adverse events, and a list of adverse events that led to the discontinuation of study drug administration, regardless of when they occurred.

The adverse events in the continuous treatment period will be tabulated in the following categories.

- Placebo group: Adverse events that will occur in the continuous treatment period.
- E0302 50mg group: Adverse events that will occur in the double-blind period and the continuous treatment period.
- Total groups: The sum of adverse events in the placebo group and E0302 50 mg group described above.

In addition to the analysis of the double-blind period described above, except for the analysis in the subpopulation, adverse events will be tabulated by the time of the first occurrence. No between-group comparison tests will be performed.

6.6.2 Laboratory test values

For physiological and hematological examinations, we will obtain the measurements at week 0 (start) and week 12 (end, baseline) of the observation period, and week 4 and 16 (end) of the treatment period and summary statistics (number of individuals, mean, SD, minimum, median, and maximum) for the changes from week 12 (end, baseline) of the observation period for each group. We will compare the changes within groups using the Wilcoxon signed-rank test and the changes between the placebo and E0302-50 mg groups using the Wilcoxon rank-sum test. We will also make a box-and-whisker diagram of the test values for each group and each period.

We will create a shift table for urinalysis at week 12 (end, baseline) of the observation period, week 4 and 16 (completion/discontinuation) of the treatment period and compare the changes within groups using the Wilcoxon signed-rank test and the changes between the placebo and E0302-50 mg groups using the Wilcoxon rank-sum test.

We will determine clinically significant abnormalities (Protocol Attachment 3: Criteria for

Confirmation of Abnormal Variations) and will determine the frequency of increase and decrease for each group during week 16 (completion/discontinuation) of the treatment period. The frequency of abnormal changes will be compared between the placebo group and the E0302 50 mg group using Fisher's direct probability test.

For the measurements of the continuous treatment period, the same analysis will be performed for the change from week 0 (start) of the observation period to each time point of the continuous treatment period.

6.6.3 Vital signs

We will perform analyses similar to those described in Section 6.6.2, "Laboratory Test Values."

6.6.4 Electrocardiogram

Measurements for heart rate, RR interval, PR interval, QRS width, QT interval, QTcB, QTcF, central reading QTcB, and central reading QTcF at baseline and at each time point of 8 or 16 weeks of the treatment period will be compared between groups using a mixed-effects model with post-dose values as the objective variable, pre-dose values as the covariate, treatment group as the fixed effect, and patient as the variable effect. Summary statistics (number of patients, mean, standard deviation, minimum, median, and maximum) of the pre- and post-dose differences in the mean values of the two measurements for each group will be obtained at baseline and at each time point of 8 or 16 weeks of the treatment period.
We will create a shift table for ECG abnormalities at week 0 (baseline) and 16 (completion/discontinuation) of the treatment period and compare its frequency within groups using the Wilcoxon signed-rank test and between the placebo and E0302-50 mg groups using the Wilcoxon rank-sum test.

6.7 Exploratory analyses

None.

6.8 The basis for setting the target number of subjects

The target number of subjects in this study was set at 64 subjects in each group and 128 subjects in both groups. The rationale for setting the target number of subjects and the history of the change are described below.

Analysis method

This study is a randomized controlled trial of two groups (placebo group and E0302 50 mg group) with the primary endpoint of the change in ALSFRS-R total score from baseline to week 16 of the treatment period, and validate the superiority of E0302 over placebo. In the MMRM model using a mixed-effects model, the change in ALSFRS-R total score from baseline to week 16 of the treatment period will be compared between the placebo group and the E0302 50 mg group. The results will be considered significant if the upper limit of the 95% confidence interval of the least-squares mean of the difference is greater than 0.

Endpoints and Estimates Used to Calculate Target Sample Size

In Phase II/III study (E0302-J081-761), an estimate of the change in total ALSFRS-R scores was estimated using data from a subset of patients who experienced 1 year or less after the onset of symptoms at the beginning of the observation period and whose total ALSFRS-R scores decreased by 1-2 points during the observation period (12 weeks). The mean ± SD change in ALSFRS-R

sum scores at 16 months were －3.2 ± 4.0 in the mecobalamin 50 mg group (26 patients in the E0302-50-mg group in this study) and －5.8 ± 5.0 in the placebo group (n = 32) (mean difference: －2.6). Based on these results, the target sample size calculation for this study assumes that the ALSFRS-R sum score for the E0302-50 mg group is －3.2, the ALSFRS-R sum score for the placebo group is －5.8, and the E0302-50 mg group outperformed the placebo group by Δ (2.6 points).
To account for the variation because of changes in ALS diagnostic criteria and participating centers compared to the Phase II/III study (E0302-J081-761), we set a more significant standard deviation (5.0) in the mecobalamin 50 mg group and the placebo group (in terms of change in ALSFRS-R total score) as the standard SD across the study.

Calculation of target sample size

A minimum of 60 individuals per group will be required to achieve a type I error probability of 2.5% or less and a power of 80% or more in a one-sided test. In addition, in view of withdrawals during the study, the target number of individuals for this study will be 64 per group, totaling 128.

Two group *t*-test of equal means (equal n's)

	1
Test significance level, α	0.025
1- or 2-sided test?	1
Group 1 mean, μ_1	−3.200
Group 2 mean, μ_2	−5.800
The difference in means, $\mu_1 - \mu_2$	2.600
Common standard deviation, σ	5.000
Effect size, $\delta = \lvert\mu_1 - \mu_2\rvert / \sigma$	0.520
Power (%)	80
n per group	60

7 Appendix

None

8 Reference material

1) Protocol 5.0 Version (April 24, 2019).
2) Statistical analysis plan 1.1 Version (February 15, 2018)

9 Citation

None

10 Revision history

Version	Draft date	Main contents	The Author(s)
1.1	2018.2.15	First edition	Tatsuo Kagimura
2.1	2020.3.24	Final analysis plan version	Tatsuo Kagimura

Issue
〒650-0047 1-5-4, Minatoshima-Minami-cho, Chuo-ku, Kobe
Medical Innovation Promotion Center
TEL: 078-303-9107 FAX: 078-303-9094
URL: http://www.tri-kobe.org