Cover Page – Statistical Analysis Plan

Study Official Title: A Phase 2 Pharmacodynamic Trial of Ezogabine on Neuronal Excitability in Amyotrophic Lateral Sclerosis

NCT #: NCT02450552

Version Date of the Document: September 27, 2018

STATISTICAL ANALYSIS PLAN (SAP)

Title	A Phase 2 Pharmacodynamic Trial of Ezogabine on Neuronal Excitability in Amyotrophic Lateral Sclerosis
Regulatory Sponsor	Brian J. Wainger, MD, PhD
Current Protocol Version	5.0
Current Protocol Date	09 Mar 2015
Statistical Analysis Plan Version	1.0
Statistical Analysis Plan Date	27 Sep 2018

SAP APPROVAL SIGNATURES

Brian J. Wainger, MD, PhD
Principal Investigator and Sponsor

Date

Eric A. Macklin, PhD
Study Biostatistician

Date

SAP REVISION HISTORY

Version	Date	Description of Changes
1.0	27 Sep 2018	Initial version

Table of Contents

1. Introduction 5
2. Study Design 5
2.1 Overview 5
2.2 Study Objectives 5
2.3 Study Populations 5
2.4 Participant Flow and Site Approval 6
2.5 Treatment Allocation 6
2.6 Treatment Administration 6
2.7 Allocation Concealment 7
2.8 Schedule of Assessments for ALS Participants 7
3. Statistical Methodology 8
3.1 General Considerations 8
3.1.1 Statistical Software 8
3.1.2 Summary Statistics 8
3.1.3 Precision 9
3.2 Study Endpoints 9
3.2.1 Efficacy Endpoints 9
3.2.2 Safety Endpoints 10
3.2.3 Tolerability Endpoint 10
3.3 Measurement Definitions 10
3.3.1 TMS Measures 10
3.3.2 TTNCS Measures 11
3.3.3 ALSFRS-R 12
3.3.4 SVC 12
3.3.5 APB Strength 12
3.3.6 Cramping 12
3.3.7 Fasciculations 12
3.3.8 AUA Symptom Score 13
3.4 Determination of Sample Size 13
3.5 Analysis Samples 13
3.6 Baseline Characterization 14
3.7 Interim Analysis 14
3.8 Efficacy Analysis 14
3.8.1 Primary Analysis of the Primary Efficacy Endpoint 14
3.8.2 Secondary Analyses of the Primary Efficacy Endpoint 15
3.8.3 Secondary Efficacy Endpoints 16
3.8.4 Subgroup Analyses 16
3.8.5 Multiplicity Adjustments 16
3.8.6 Missing Data 17
3.9 Safety Analysis 17
3.9.1 Treatment-emergent Adverse Events 17
3.9.2 Suicidality 17
3.9.3 Safety Labs 17
3.9.4 Vital Signs 17
3.9.5 Additional Continuous Safety Outcomes 18
3.9.6 Post-study Mortality 18
3.10 Other Analyses 18
3.10.1 Participant Disposition 18
3.10.2 Study Drug Compliance and Tolerance 18
3.10.3 Prior and Concomitant Medication Use 18
4. References 18

1. Introduction

This Statistical Analysis Plan (SAP) defines the outcome measures and analysis samples and specifies the planned analyses of data from Dr. Brian Wainger's Ezogabine trial. This SAP is specific to analysis of data related to the ALS participants in the randomized intervention trial. The SAP supplements the clinical protocol. Please refer to the clinical protocol for details on the rationale for the intervention, eligibility criteria, conduct of the trial, clinical assessments and the timing of their use in the trial, definitions and reporting of adverse events, data management conventions, and regulatory oversight and compliance procedures. In case of discrepancies between the SAP and the clinical protocol concerning matters of analysis, the SAP is authoritative. On all other matters, the clinical protocol is authoritative.

2. Study Design

2.1 Overview

This is a multicenter, double-blind, placebo-controlled, three-arm, parallel-group, phase 2 randomized controlled trial to examine the effect of two dosages of ezogabine on neuronal excitability in ALS patients. The trial also enrolls unmatched and age-matched healthy control subjects to evaluate disease-dependent differences in baseline measures and longitudinal changes in neuronal excitability. Trial participation for ALS participants includes a screening visit, randomization to 600 mg/day or 900 mg/day ezogabine or placebo, 10 weeks of follow-up on study drug, and 4 weeks of follow-up after last dose of study drug. Trial participation for unmatched healthy control (UHC) participants includes a screening visit and up to 4 additional neurophysiological testing visits with one or two neurophysiologists over a span of up to 6 weeks. Trial participation for age-matched healthy control (MHC) participants includes a screening visit and two follow-up visits over 12 weeks. The trial is registered at Clinicaltrials.gov as study NCT02450552 (see https://clinicaltrials.gov/ct2/show/NCT02450552).

2.2 Study Objectives

The primary efficacy objective of the trial is to determine whether ezogabine reduces neurophysiological excitability indices in ALS participants as assessed by transcranial magnetic stimulation (TMS) and threshold tracking axonal nerve conduction studies (TTNCS). Secondary efficacy objectives include evaluating the effect of ezogabine on measures of clinical progression, including functional scores, vital capacity, and muscle strength, and on measures of muscle cramping and fasciculations.

The safety objectives of the trial are to assess the safety and tolerability of ezogabine by assessing treatment-emergent adverse events (TEAEs), vital signs, laboratory assessments, electrocardiography (ECG), urologic symptoms, ophthalmologic symptoms, and suicidality.

2.3 Study Populations

Individuals eligible for trial participation as an ALS participant are men or women at least 18 years old who meet the El Escorial criteria of possible, laboratory-supported probable, probable, or definite criteria for a diagnosis of ALS and have a slow vital capacity (SVC) at least 50% of that expected based on age, sex, and height. Individuals eligible for study participation as an UHC participant are men or women at least 18 years old who do not have a history or family

history of possible motor neuron disease. Individuals eligible for study participation as a MHC participant are men or women who do not have a history or family history of possible motor neuron disease and are within 10 years of age of an enrolled ALS participant. Detailed inclusion and exclusion criteria are specified in the clinical protocol.

Participants will be recruited from a total of 12 clinical sites located through the US.

2.4 Participant Flow and Site Approval

After providing informed consent, approximately 5 UHC participants will be screened for eligibility at each site, including assessment of neurophysiology by TMS and TTNCS by one or two neurophysiologists. Each UHC participant will have up to 4 additional neurophysiological testing visits over a span of up to 6 weeks.

After completing TMS and TTNCS assessments of the first UHC participant, sites must receive approval from the Coordination Center before continuing enrollment of UHC participants. Approval will be based on review of TMS and TTNCS data by the respective central reading centers. After completing enrollment of approximately 5 UHC participants, sites must receive approval from the Coordination Center before initiating recruitment of ALS participants.

After a site is activated for recruitment of ALS participants and after each ALS patient provides informed consent and is determined eligible, approximately 120 ALS participants will be randomized to receive ezogabine 600 mg/day, ezogabine 900 mg/day, or placebo in a 1:1:1 ratio, stratified by clinical site. The first dose of stdy drug is administered at the baseline visit. ALS participants are exposed to study drug for 10 weeks, from baseline to a week 10 telephone call. ALS participants complete in-clinic evaluations at weeks 4, 6, 8, and 12 and complete follow-up by telephone at weeks 2, 10, and 14 to assess for TEAEs and update concomitant medication logs. Participants who discontinue study drug should remain on study following the normal schedule of assessments. Patients withdrawing consent are asked to delay withdrawing consent until after they return for a Final Safety Visit approximately 2 weeks after their last dose of study drug and after they complete a final telephone call 4 weeks after their last dose of study drug.

Detailed descriptions of study procedures and timing are specified in the clinical protocol.

2.5 Treatment Allocation

Prior to the baseline visit, eligible ALS participants are randomly allocated in equal proportions to one of three treatment groups, ezogabine 600 mg/day, ezogabine 900 mg/day, or placebo, according to a permuted-block randomization schedule, stratified by site. The randomization schedule was prepared by computer program by the unblinded study statistician.

2.6 Treatment Administration

Study drug will be orally self-administered in film-coated immediate-release tablets of a variety of strengths of ezogabine or matching placebo. In the 600 mg/day arm, ezogabine is titrated as follows: 100 mg TID for week 1, 150 mg TID for week 2, 200 mg TID for weeks 3 through 8, 200 mg BID for week 9, and 200 mg QD for week 10. In the 900 mg/day arm, ezogabine is titrated as follows: 100 mg TID for week 1, 150 mg TID for week 2, 200 mg TID for week 3, 250 mg TID for week 4, 300 mg TID for weeks 5 through 8, 200 mg BID for week 9, and 200 mg QD for week 10. In the placebo arm, participants receive a matching pseudo-titration of placebo tablets matched for size, color, and presentation.

2.7 Allocation Concealment

The randomization schedule is known only by the unblinded study statistician who generated the schedule and by the study drug distributor. Concealment of the true treatment allocation of specific participants is achieved by use of anonymous subject identifiers to link participants with specific study drug kits and by use of matching active and placebo tablets and titration schedules. Clinical members of the Steering Committee, site investigators and other site staff, clinical coordination and data management staff, the medical monitor, and all participants are blinded to participant treatment allocations. The Data and Safety Monitoring Board (DSMB) members are provided treatment-specific information in order to monitor the trial but such information is masked by use of coded values to identify the treatment groups. The DSMB may request the true treatment identities.

2.8 Schedule of Assessments for ALS Participants

Activity	Screening	Baseline[2]	Week 2	Week 4 [3]	Week 6[3]	Week 8[3]	Week 10	Week 12	Week 14
	CLINIC	CLINIC	CALL	CLINIC	CLINIC	CLINIC	CALL	CLINIC	CALL
	Day -21 to -1	Day 0	Day 14 ± 3	Day 28 ± 3	Day 42 ± 3	Day 56 ± 3	Day 70 ± 3	Day 84 ± 3	Day 98 +5
Written Informed Consent	X								
Inclusion/Exclusion Review	X								
Medical History	X								
Demographics	X								
ALS Diagnosis History	X								
Concomitant Medication Review	X	X	X	X	X	X	X	X	X
Adverse Event Review[4]	X	X	X	X	X	X	X	X	X
Physical Examination	X							X	
Neurological Exam	X							X	
Manual Muscle Testing (MMT)	X							X	
Vital Signs[5]	X	X		X	X	X		X	
12-Lead Electrocardiogram (ECG)	X	X[15]			X			X	
Height & Weight[6]	X	X		X	X	X		X	
Safety Labs[7]	X			X	X	X		X	
Urological Symptom Score	X			X		X		X	
ALSFRS-R	X	X		X	X	X		X	
Slow Vital Capacity (SVC)	X	X		X	X	X		X	
Hand Held Dynamometry (APB only)	X	X		X	X	X		X	
Assign Global Unique Identifier (GUID)	X								
Blindedness Questionnaire and Exit Survey								X	
Suicide Assessment (C-SSRS)		X		X	X	X		X	
Edinburgh Handedness Inventory – short form	X								
Blood collection for iPSC Lines				X[8]	(X[8])	(X[8])			
Neurophysiological Testing (TMS/NCS)	X	X		(X[17])	X	X		(X[17])	
Ophthalmologic Examination[9]		X						X	
Lumbar Puncture (Optional)					X[16]				
Pharmacokinetics (PK)/Biobanking					X				

Activity	Screening CLINIC	Baseline[2] CLINIC	Week 2 CALL	Week 4 [3] CLINIC	Week 6[3] CLINIC	Week 8[3] CLINIC	Week 10 CALL	Week 12 CLINIC	Week 14 CALL
	Day -21 to -1	Day 0	Day 14 ± 3	Day 28 ± 3	Day 42 ± 3	Day 56 ± 3	Day 70 ± 3	Day 84 ± 3	Day 98 +5
Telephone Call[10]			X				X		X
Randomization[11]		X							
Administer/Dispense Study Drug		X[12]		X					
Drug Accountability/ Compliance		X		X[13]	X[13]	X[13]		X[13]	
Review of Muscle Symptoms Diary[14]		X		X	X	X		X	
Fasciculation Assessment Form		X		X	X	X		X	

[1] All clinic visits after the Baseline Visit and before Week 14 will have a ± 3 day window.

[2] Screening procedures must be completed prior to the Baseline Visit, which will be scheduled up to 21 days after the Screening Visit.

[3] Study staff should document the time the last dose of ezogabine was administered.

[4]Adverse events that occur AFTER signing the informed consent form will be recorded.

[5] Vital signs include systolic and diastolic pressure in mm Hg, respiratory rate/minute, heart rate/minute and temperature.

[6] Height measured at Screening Visit only.

[7] Labs include Hematology (CBC with differential), Complete Chemistry Panel, Liver Function Tests, Urinalysis and serum pregnancy test (WOCBP, screening visit only).

[8] Blood for iPSC lines can be collected at either week 4, 6 or 8.

[9] Ophthalmologic Examinations may be scheduled separately, but within 14 days prior to Baseline Visit and within 14 days of Week 12 Visit.

[10] Subjects will be called at the designated time points, as a reminder to escalate up or taper off study drug, when appropriate.

[11] Randomization should occur at or prior to the Baseline Visit.

[12] Administer first dose of study drug AFTER all Baseline Visit procedures are completed (with the exception of ECG testing). Subjects to take study drug until the Week 10 call.

[13] Review dosing diary

[14] Muscle symptoms will be completed every day, starting at the Baseline visit, with a Muscle Symptom Diary and should be reviewed by study staff at every visit after Baseline.

[15] ECG testing should be performed 3 hours post initial study drug administration.

[16] Lumbar puncture is optional

[17] Only participant enrolled prior to protocol version 4 were scheduled to complete TMS and TTNCS testing at the week 4 and 12 visits.

3. Statistical Methodology

3.1 General Considerations

3.1.1 Statistical Software

All statistical analyses will be performed using SAS (SAS Institute, NC, USA) or R (R Foundation for Statistical Computing, Vienna, Austria).

3.1.2 Summary Statistics

Data will be summarized with respect to disposition, demographics, pre-treatment characteristics, safety outcomes, tolerability, and efficacy outcomes. Summary statistics for continuous variables will include the number of subjects, the mean, median, standard deviation, and range. For categorical data, summaries will include counts and percentages.

3.1.3 Precision

Results will generally be reported to 3 significant figures. Percentages will generally be reported to 0.1 percentage points. P-values will be reported to two digits when greater than or equal to 0.095, to three digits when greater than or equal to 0.00095 and less than 0.095, and as <0.001 for all smaller values.

3.2 Study Endpoints

3.2.1 Efficacy Endpoints

The primary efficacy endpoint is the change from pre-treatment to the average of the week 6 and week 8 visits in paired-pulse short-interval intracortical inhibition (SICI) assessed by TMS.

The key secondary efficacy endpoint is the change from pre-treatment to the average of the week 6 and week 8 visits in resting motor evoked potential (MEP) threshold (RMT) assessed by TMS.

The other secondary efficacy endpoints are the change from pre-treatment to the average of the week 6 and week 8 visits in the following measures.

TMS measures:

- MEP amplitude at 120% of RMT,
- Intracortical facilitation (ICF), and
- Cortical silent period (CSP).

TTNCS measures:

- Strength duration time-constant (SDTC),
- Depolarizing and hyperpolarizing electrotonus at 90 to 100 ms (TEd(90-100 ms) and TEh(90-100 ms)),
- Recovery cycle measures of relative refractory period (RRP), superexcitability, and subexcitability, and
- Compound motor action potential (CMAP).

Clinical measures:

- ALS Functional Rating Scale-Revised (ALSFRS-R) total score,
- Slow vital capacity (SVC),
- Abductor pollicis brevis (APB) muscle strength by hand-held dynamometry (HHD),
- Muscle cramping frequency and pain, and
- Fasciculation occurrence and interference with daily life.

Exploratory efficacy measures include: maximum amplitude, minimum amplitude, midpoint of activation, and slope at the midpoint of activation from input/output curves; rheobase, TEd(40-60 ms), TEh(10-20 ms), TEd(peak), subexcitability after second pre-pulse, and stimulus for 50% maximum response assessed by TTNCS; MEP amplitude at 120% of RMT divided by CMAP.

3.2.2 Safety Endpoints

The following safety endpoints will be evaluated:

- Proportion of participants experiencing and number of unique events of each type of TEAE and serious TEAE classified by MedDRA system organ class and preferred term,
- Proportion of participants experiencing and number of unique events of TEAEs classified by seriousness, severity, relatedness to study drug, relatedness to TMS and TTNCS procedures, action taken with study drug, action taken with TMS/TTNCS, and outcome, summarized across all MedDRA terms,
- Instances of treatment-emergent laboratory abnormalities classified by assay, whether the abnormality is below or above the normal range, and extend of deviation from normal,
- Mean change from baseline in weight
- Proportion of participants experiencing treatment-emergent clinically significant ECG abnormalities
- Proportion of participants experiencing treatment-emergent ophthalmologic complication,
- Mean change in AUA Symptom Score and proportion of participants with mild, moderate, or severe symptoms by visit,
- Proportion of participants reporting any treatment-emergent suicidal ideation
- Maximum post-baseline suicidal ideation planning, frequency, duration, controllability, deterrents, and reasons
- Proportion of participants reporting any post-baseline suicidal behavior, actual attempts, non-suicidal self-injurious behavior, interrupted attempts, aborted attempts, preparatory acts or behaviors, or completed suicide
- Maximum post-baseline suicidal behavior lethality and potential lethality

Reported proportions will use as their denominator all participants in the Safety and Tolerability sample (see Section 3.5 below).

3.2.3 Tolerability Endpoint

Participants will be judged tolerant of study drug if they reached their target dose and remain on study drug until planned discontinuation. Tolerability will be summarized as the proportion of participants in a treatment group who are tolerant of study drug.

3.3 Measurement Definitions

3.3.1 TMS Measures

Transcranial magnetic stimulation (TMS) is a neurophysiologic test for assessing upper motor neuron function. A magnetic coil is positioned on the head above the motor cortex at a location that causes excitation of the abductor pollicis brevis (APB) muscle in a participant's dominant hand. High-frequency sampling of motor evoked potentials (MEP) in the APB is used to record neurophysiologic response to specific magnetic field strengths and sequences. The MEP traces are used to calculate a range of derived measures. MEPs that were below the limit of detect and recorded as zero mV were replaced by 0.005 mV.

- Resting motor threshold (RMT): RMT is the magnetic field strength, measured as a percentage of the maximum stimulator output, that produces at least a 0.05 mV response in at least 5 of 10 consecutive trials.
- MEP amplitude: MEP amplitude is defined as the response when a stimulus equal to 120% of RMT is administered. MEP amplitude is calculated as the geometric mean of replicate estimates.
- Short-interval intracortical inhibition (SICI): Paired-pulse SICI is defined as the ratio of the response after a conditioning pulse equal to 80% of RMT is administered 3 ms prior to the signaling pulse divided by MEP amplitude. SICI is calculated as the geometric mean of replicate estimates.
- Intracortical facilitation (ICF): Paired-pulse ICF is defined as the ratio of the response after a conditioning pulse equal to 80% of RMT is administered 15 ms prior to the signaling pulse divided by MEP amplitude. ICF is calculated as the geometric mean of replicate estimates.
- Cortical silent period (CSP): CSP is the suppression of voluntary muscle contraction elicited by stimulation equal to 120% of RMT. CSP duration is measured from the time of the muscle activity suppression to return of muscle activity.
- Input/Output curve parameters: MEPs elicited by stimuli ranging in strength from 60% of RMT to 150% of RMT follow a sigmoidal curve with upper and lower thresholds and a smooth, symmetric transition between the thresholds. A four-parameter logistic function fit to the input/output curve after log-transformation of output potentials yields estimates of maximum amplitude, minimum amplitude, midpoint of activation, and slope at the midpoint of activation.
- TMS quality: Good vs. poor quality TMS assessments will be evaluated as a measure of electrophysiologic response. Good quality will be defined as a central reader's evaluation that the TMS records were of good quality overall and that individual measurements of RMT, input/output curves, and CSP were of good quality.

TMS tracings were centrally reviewed and processed. Central readers indicated the reliability of the overall TMS assessment and of individual measures using a set of quality flags. Only TMS assessments that were evaluated as good quality overall and for measurement of RMT, input/output curves, and CSP are used for primary analysis. CSP times equal to zero were considered low quality whether flagged as such or not.

3.3.2 TTNCS Measures

Threshold-tracking nerve conduction studies (TTNCS) is a neurophysiologic test for assessing lower motor neuron function. The following measures were obtained using standard procedures (Kiernan et al. 2000): strength-duration measures of strength duration time-constant (SDTC) and rheobase; threshold electrotonus measures of depolarizing threshold electrotonus at 40 to 60 ms [TEd (40-60 ms)], 90 to 100 ms [TEd (90-100 ms)], and at peak [TEd (peak)] and hyperpolarizing threshold electrotonus at 10 to 20 ms [TEh (10-20 ms)] and 90 to 100 ms [TEh (90-100 ms)]; recovery cycle measures of RRP, superexcitability, subexcitability after first pre-pulse, and subexcitability after second pre-pulse; peak amplitude measures of compound motor evoked potential (CMAP) and stimulus for 50% maximum response.

TTNCS tracings were centrally reviewed and processed. A central reader indicated the reliability of the overall TTNCS assessment and of individual measures using a set of quality flags. For primary analysis, all measures were excluded if the overall quality flag indicated that the full assessment was unusable. Otherwise, acceptable quality for primary analysis was separately indicated for peak amplitude (CMAP, stimulus for 50% maximum response), strength-duration measures (SDTC, rheobase), threshold electrotonus (TEd (40-60 ms), TEd (90-100 ms), TEd (peak), TEh (10-20 ms), TEh (90-100 ms)), RRP, superexcitability, and subexcitability (after first and second pre-pulse).

3.3.3 ALSFRS-R

The ALSFRS-R (Cedarbaum et al. 1999) is a 12-item clinician-completed interview scale for assessing participants' function in four domains: bulbar, fine motor, gross motor, and respiratory. Each item is scored from 0 to 4 with higher scores indicating greater function. The ALSFRS-R total score is the sum of all items (range 0 to 48).

3.3.4 SVC

Slow vital capacity (SVC) is the maximum volume of air that can be slowly exhaled after slow, maximal inhalation. Trained technicians coach participants through 3 to 5 maneuvers using a spirometer. The maximum volume expired is converted to percent of predicted normal. Normal values are calculated based on sex, age, and height using equations published by Knudsen et al. (1983).

3.3.5 APB Strength

Bilateral abductor pollicis brevis (APB) muscle strength will be measured quantitatively by hand-held dynamometry (HHD) and assessed subjectively by manual muscle testing (MMT). MMT is a single 8-level item with scores ranging from 0 if no contraction can be achieved and 5 if the muscle exerts normal strength. Scores of 4 are subdivided into three levels ranging from 4- if active movement is achieved against the force of gravity and slight resistance to 4+ if active movement is achieved against the force of gravity and strong resistance.

3.3.6 Cramping

Frequency of muscle cramping and maximum pain from muscle cramping is collected at baseline by retrospective report of the prior 7 and 30-day intervals. Prospective muscle cramping and pain from muscle cramping is recorded starting at baseline by participants in daily diaries and then summarized as weekly summaries of total number of cramps experienced and maximum and average daily maximum pain from muscle cramping.

3.3.7 Fasciculations

Data on muscle fasciculations are captured at baseline and each post-baseline clinic visit by retrospective report of the prior 14-day interval. Data are captured on the experience of any fasciculations, the average duration of fasciculations, the association of fasciculations with exertion or times of rest, the degree to which fasciculations interfere with daily activities, and interference of fasciculations with sleep. Prospective data on fasciculations are recorded starting at baseline by participants in daily diaries and then summarized as weekly summaries indicating the number of days with no fasciculations, with fasciculations after exertion that did not substantially interfere with daily activities, with fasciculations at rest or after exertion that did

not substantially interfere with daily activities, and fasciculations that substantially interfered with daily activities.

3.3.8 AUA Symptom Score

The American Urological Association (AUA) Symptom Index (Barry et al. 2017) is a 7-item instrument addressing urinary frequency, nocturia, weak urinary stream, hesitancy, intermittence, incomplete emptying, and urgency. Each item is scored from 0 to 5 with higher scores indicating greater symptom severity. The AUA Symptom Score is the sum of all items (range 0 to 35). AUA Symptom Scores from 0 through 7 are classified as mild symptoms, scores from 8 through 19 are classified as moderate symptoms, and scores of 20 through 35 are classified as severe symptoms.

3.4 Determination of Sample Size

Caramia et al. (2000) report mean (SD) paired pulse SICI at 3 ms among 16 ALS patients of 125.7 (25.7) percent of MEP control size. Stefan et al. (2001) report paired pulse SICI at 3 ms among 13 ALS patients of 96.9 (35.9) percent of MEP control size. Assuming a mean (SD) equal to 100 (30) percent of MEP and using a two-sided alpha of 0.027 based on Dunnett's correction for comparing two treatments to placebo, 40 patients in each group and up to 25% loss to follow-up would give 80% power to detect a 25% difference between placebo and each active treatment based on a simple two-sample t-test. This accommodates the loss of power due to early stopping for futility and is conservative relative to our proposed shared-baseline mixed model analysis. For reference, Caramia et al. (2000) report an effect of three-week exposure to gabapentin equal to an 80% reduction of the paired pulse amplitude in untreated patients, and Stefan et al. (2001) report an effect of 5 or more day exposure to riluzole equal to a 35% reduction of the conditioned stimulus amplitude.

3.5 Analysis Samples

The following analysis samples will be used for testing safety, tolerability, and efficacy endpoints:

- Safety and Tolerability (ST) Sample: ALS participants who are eligible, randomized, and take at least one dose of study drug, classified according to the actual treatment received. Observations made after premature permanent discontinuation of study drug are included in this sample, should such participants remain on study.
- Efficacy Modified Intent-to-treat (mITT) Sample: ALS participants who are eligible, randomized, and take at least one dose of study drug, classified according to their randomized treatment assignment. Observations made after premature permanent discontinuation of study drug are included in this sample, should such participants remain on study.
- Efficacy Per-protocol (PP) Sample: ALS participants who are eligible, randomized, and take at least one dose of study drug, classified according to the actual treatment received. Observations made after premature permanent discontinuation of study drug (whether due to intolerance, early termination, loss to follow-up, or death) are excluded from this sample.

- Efficacy Low RMT (loRMT) Sample: ALS participants and observations in the PP sample with RMT less than 67% of maximum stimulator output. This criterion ensures that the full TMS input/output curve could be conducted.
- Efficacy All Quality (AllQual) Sample: ALS participants and observations in the PP sample with no restriction that TMS quality is good.

3.6 Baseline Characterization

Each analysis sample will be summarized overall and by treatment group for the following characteristics: randomization site; age, sex, race, ethnicity, El Escorial diagnosis, years since ALS symptom onset, delay between symptom onset and ALS diagnosis, site of ALS symptom onset, use of riluzole at baseline, ALSFRS-R total score, SVC, APB muscle strength by HHD on the dominant hand, TMS parameters, TTNCS parameters, cramping frequency, cramping pain, fasciculation intensity, fasciculation interference with daily living, fasciculation interference with sleep, and AUA Symptom Score. The magnitude of differences between treatment groups will be summarized using Fisher's exact tests for nominal variables and two-sample t-tests for approximately continuous variables.

3.7 Interim Analysis

An interim analysis for futility was planned after complete follow-up data were available from the 45th ALS participant; however, because the final randomization was scheduled less than three months after this timepoint due to slower than anticipated enrollment, the interim analysis was not conducted, with DSMB approval. The independent DSMB reviewed safety data quarterly and could request specific analyses at any time.

3.8 Efficacy Analysis

3.8.1 Primary Analysis of the Primary Efficacy Endpoint

The primary efficacy analysis will be performed on SICI of the APB after log-transformation in the mITT sample of ALS participants and will use a shared-baseline mixed model with fixed effects for visit (6 levels: Screening, Baseline, and Weeks 4, 6, 8, and 12), treatment group (3 levels: placebo and 600 and 900 mg/day ezogabine) x post-baseline visit (4 levels) and random site-specific intercepts and random participant-specific intercepts and slopes with unstructured covariance. The interaction between fixed effects for treatment group and visit will be restricted to post-baseline visits by including a numeric indicator variable (0 pre-treatment, 1 post-treatment) in the interaction. Use of a shared baseline reflects the true state of the population sampled prior to randomization and has the advantage of adjusting for any chance differences at baseline in a manner similar to analysis of covariance (Liang and Zeger, 2000) with potentially greater efficiency with more than one follow-up visit and drop-out. Random participant-specific slopes will be estimated per month from treatment initiation with all pre-treatment observations assigned a value of 0 months from treatment initiation. Data from Final Safety or Early Termination visits will be assigned to the next scheduled post-baseline visit. The following SAS code specifies the model:

```
proc mixed data=xxx method=reml;
 class site id trtrnd visit;
 model Value = visit post*trtrnd*visit;
 random intercept / subject=site type=vc;
 random intercept month / subject=id(site) type=un;
```

Treatment-dependent differences in the change from pre-treatment to the average of the week 6 and 8 visits on SICI will be estimated by a linear contrast and tested using a two-tailed Wald-test at alpha = 0.027. The following SAS code specifies the linear contrasts used to obtain the two co-primary estimands (with the sort order for treatment group being placebo, 600 mg/day ezogabine, and 900 mg/day ezogabine, and the sort order for visit being chronological):

```
estimate "2|600vsPlb|dAvg 6&8"
      post*trtrnd*visit 1 1 0 -1 -1 0  -1 -1 0 1 1 0  0 0 0 0 0 0 / cl divisor=2;
estimate "2|900vsPlb|dAvg 6&8"
      post*trtrnd*visit 1 1 0 -1 -1 0  0 0 0 0 0 0  -1 -1 0 1 1 0 / cl divisor=2;
```

Estimates and their 95% confidence bounds will be back-transformed for reporting as proportional changes. A significant mean reduction in SICI from pre-treatment to the average of weeks 6 and 8 among participants randomized to either 600 or 900 mg/day ezogabine relative to mean changes among participants randomized to placebo would be considered evidence that ezogabine is effective in reducing cortical hyperexcitability.

3.8.2 Secondary Analyses of the Primary Efficacy Endpoint

Several secondary analyses will be investigated to assess sensitivity of our estimates of treatment effect to alternative modeling assumptions.

- Active vs. placebo: A secondary estimand of the efficacy of ezogabine will compare both 600 mg/day and 900 mg/day treatment arms vs. placebo, tested at alpha = 0.05. The following SAS code specifies the linear contrast for this comparison:

  ```
  estimate "2|ActvsPlb|dAvg 6&8"
        post*trtrnd*visit 2 2 0 -2 -2 0  -1 -1 0 1 1 0  -1 -1 0 1 1 0 / cl divisor=4;
  ```

- Random-slopes model: The visit-specific fixed effects will be replaced with a continuous predictor of month since treatment initiation. Data from Final Safety and Early Termination visits will be placed at their observed follow-up time. Only assessments completed through Week 10, the end of planned study drug exposure, will be included. The following SAS code specifies the model and the linear contrasts used to estimate treatment-dependent differences in rate of change per month:

  ```
  proc mixed data=xxx method=reml;
   where visit in ("Screening","Baseline","Week 04","Week 06","Week 08","Week 10");
   class site id trtrnd;
   model Value = month trtrnd*month;
   random intercept / subject=site type=vc;
   random intercept month / subject=id(site) type=un;
   estimate "2|600vsPbl|Slope/mn" intercept 0 month 0 trtrnd*month -1 1 0 / cl;
   estimate "2|900vsPbl|Slope/mn" intercept 0 month 0 trtrnd*month -1 0 1 / cl;
   estimate "2|ActvsPbl|Slope/mn" intercept 0 month 0 trtrnd*month -2 1 1 / cl divisor=2;
  ```

- Shared-baseline mixed model with pMI for missing data: The shared-baseline mixed model with fixed effects for visits described above for the primary analysis will be applied to datasets generated by placebo multiple imputation (Ayele et al. 2014). Fifty imputed datasets will be generated for a given endpoint.

- PP sample: The primary efficacy model and the secondary random-slopes model and active vs. placebo estimand will be applied to the PP sample.

- loRMT sample: The primary efficacy model and the secondary random-slopes model and active vs. placebo estimand will be applied to the loRMT sample.

- AllQual sample: The primary efficacy model and the secondary random-slopes model and active vs. placebo estimand will be applied to the AllQual sample.

3.8.3 Secondary Efficacy Endpoints

The key secondary efficacy endpoint of RMT will be analyzed using the same analysis samples and models as specified for the primary efficacy endpoint in Sections 3.8.1 and 3.8.2. Continuous non-key secondary efficacy endpoints and exploratory efficacy endpoints measuring electrophysiology and clinical progression will all be analyzed using the mITT sample and the primary efficacy analysis described in Section 3.8.1. Variables that are strongly right-skewed (e.g., MEP amplitude, ICF) will be log-transformed prior to analysis, and estimates will be back-transformed for reporting.

The binary indicator of good quality TMS recordings will be analyzed in an equivalent generalized mixed model treating the indicator variable as a Bernoulli random variable with logit link.

Weekly cramping frequency data will be analyzed in similar generalized mixed models treating the number of cramps experienced in a week as a negative binomial random variable with log link and with the log of the number of days at risk for cramps as an offset. Similarly, weekly number of days with fasciculations will be analyzed in similar generalized mixed models treating the number of days with fasciculations in a week as a binomial random variable with logit link. Because diary data are only collected after baseline, fixed effects in the model will include treatment group, post-baseline week, and baseline retrospective report of number of cramps experienced in the prior 30-day interval or the experience of fasciculations in the prior 14-day interval.

3.8.4 Subgroup Analyses

Differences in treatment efficacy will be explored in several pre-defined subgroups: baseline SICI (median split), baseline MEP amplitude (median split), and baseline riluzole use (yes vs. no). Subgroup specific estimates will be obtained by including subgroup, subgroup × visit, and subgroup × treatment × post-baseline visit terms to models for primary and key secondary efficacy outcomes based on the mITT sample.

3.8.5 Multiplicity Adjustments

The effect of each dose of ezogabine on the primary efficacy outcome in the mITT sample will be tested at two-tailed $p < 0.027$ following Dunnett's correction for comparing two active treatments to placebo. If this primary analysis identifies a significant benefit from ezogabine treatment, then testing the key secondary efficacy outcome at $p < 0.027$ maintains an overall type I error rate at 5% under a closed testing sequential analysis. If significance of the key secondary outcome alone is accepted as evidence of a therapeutic benefit from ezogabine when the primary analysis is not significant, then the overall type I error rate will be limited to 10% or less (depending on the correlation between the primary and key secondary outcomes). Results from analysis of non-key secondary efficacy endpoints, exploratory efficacy endpoints, and subgroup analyses will report nominal, comparison-wise p-values, recognizing that the totality of results will be evaluated in judging the potential efficacy of ezogabine.

3.8.6 Missing Data

Baseline values for efficacy endpoints will be determined from the last non-missing data collected prior to the first dose of study medication. The planned mixed model yields estimates that are unbiased conditional on the observed scores under a missing at random assumption. In addition, a secondary analysis of the primary endpoint will use placebo-based multiple imputation of missing data. Additional sensitivity analyses may be pursued to impute missing values or otherwise construct models for unobserved outcomes if more than 20% of participants are missing follow-up data for any reason.

3.9 Safety Analysis

3.9.1 Treatment-emergent Adverse Events

The incidence of TEAEs will be summarized by the number of events of a given classification experienced by participants in each treatment group and by the number and proportion of participants experiencing such an event in each treatment group in the ST sample. TEAEs will be summarized in aggregate across all MedDRA terms and separately by MedDRA system organ class and preferred term.

Aggregate summaries of TEAE grade will include characteristics of: (a) seriousness, (b) severity, (c) relatedness to study drug, (d) relatedness to TMS, (e) relatedness to TTNCS, (f) action taken with study drug, (g) action taken with study procedure, and (h) outcome. For each level of a given TEAE characteristic, summaries will include the number of events of a given classification and by the number and proportion of participants for which that level of a characteristic was the worst they experienced (treating any unknown characteristic as not worst).

3.9.2 Suicidality

The proportion of participants who report any post-baseline suicidal ideation or any post-baseline suicidal behavior will be summarized by treatment group. Suicidal behaviors will include: actual attempts, non-suicidal self-injurious behavior, interrupted attempts, aborted attempts, preparatory acts or behaviors, and completed suicide. The most severe ideation, maximal frequency, maximal duration, minimal controllability, minimal deterrents, maximal reasons for ideation, maximal actual lethality or medical damage, and maximal potential lethality will be summarized as means, standard deviations, medians, and ranges by treatment group.

3.9.3 Safety Labs

The absolute level and the absolute change from baseline for each safety laboratory assay will be summarized as means, standard deviations, medians, and ranges at each visit by treatment group. The proportion of participants with safety lab levels below the lower limit of normal or above the upper limit of normal will be summarized by treatment group by visit and at any post-baseline visit.

3.9.4 Vital Signs

The absolute level and the absolute change from baseline for vital signs will be summarized as means, standard deviations, medians, and ranges at each visit by treatment group.

3.9.5 Additional Continuous Safety Outcomes

The additional continuous safety outcomes of weight, ECG parameters, AUA Symptom Score and their absolute change from baseline will be summarized as means, standard deviations, medians, and ranges at each visit by treatment group.

3.9.6 Post-study Mortality

Long-term survival will be summarized for each treatment group by Kaplan-Meier product-limit estimates and compared between treatments by log-rank test.

3.10 Other Analyses

3.10.1 Participant Disposition

The number of patients who were screened, randomized, completed scheduled follow up, and prematurely withdrew study participation will be summarized overall and by treatment group. Reasons for screen failure and for withdrawal from study will be presented.

3.10.2 Study Drug Compliance and Tolerance

The number of days of exposure to study drug will be summarized by treatment group. Compliance with study drug will be calculated as the number of doses taken divided by the scheduled number of doses taken prior to permanent discontinuation, expressed as a percentage. Time to discontinuation of study drug will be estimated using Kaplan-Meier product-limit estimates. Treatment-dependent differences will be tested by log-rank test.

3.10.3 Prior and Concomitant Medication Use

Concomitant medications taken during the study period will be listed for each patient, coded using the World Health Organization Drug Dictionary Enhanced. The percentage of patients taking each class of medications will be summarized overall and by treatment group.

4. References

Ayele BT, Lipkovich I, Molenberghs G, Mallinckrodt CH. A multiple-imputation-based approach to sensitivity analyses and effectiveness assessments in longitudinal clinical trials. J Biopharm Stat. 2014;24(2):211-28.

Barry MJ, Fowler FJ Jr, O'leary MP, Bruskewitz RC, Holtgrewe HL, Mebust WK, Cockett AT; Measurement Committee of the American Urological Association. The American Urological Association Symptom Index for Benign Prostatic Hyperplasia. J Urol. 2017 Feb;197(2S):S189-S197.

Cedarbaum JM, Stambler N, Malta E, Fuller C, Hilt D, Thurmond B, Nakanishi A. The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function. BDNF ALS Study Group (Phase III). J Neurol Sci. 1999 Oct 31;169(1-2):13-21.

Kiernan MC1, Burke D, Andersen KV, Bostock H. Multiple measures of axonal excitability: a new approach in clinical testing. Muscle Nerve. 2000 Mar;23(3):399-409.

Knudson RJ, Lebowitz MD, Holberg CJ, Burrows B. Changes in the normal maximal expiratory flow-volume curve with growth and aging. Am Rev Respir Dis. 1983 Jun;127(6):725-34.

Liang KY, Zeger S. Longitudinal data analysis of continuous and discrete responses for pre–post designs. Sankhyā Ser B. 2000 Apr;62(1):134–148.