Developing Neuronal KCNQ Channel Modulators for Mood Disorders
PI: Dr. James Murrough
NCT03043560
Document Date: June 2018

STATISTICAL ANALYSIS PLAN

DEVELOPING NEURONAL KCNQ CHANNEL MODULATORS FOR MOOD DISORDERS

Sponsored By NIMH

Clinical Coordinating Center:
Icahn School of Medicine at Mount Sinai
New York

Data Coordination and Statistical Center:
InCHOIR
Icahn School of Medicine at Mount Sinai
New York

June 2018

CONFIDENTIAL

Table of Contents

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

ACC	Anterior cingulate cortex
ACIPS	Anticipatory and Consummatory Interpersonal Pleasure Scale
AE	Adverse event
CGI	Clinical Global Impression Scale
C-SSRS	Columbia-Suicide Severity Rating Scale
DS	Dorsal striatum
fMRI	Functional magnetic resonance imaging
IFT	Incentive flanker task
ITT	Intent-to-treat
IWRS	Interactive web response system
LOCF	Last observation carried forward
MADRS	Montgomery-Asberg Depression Rating Scale
MAR	Missing at random
MASQ	Mood and Anxiety Symptoms Questionnaire
MedDRA	Medical Dictionary for Regulatory Activities
MID	Monetary incentive delay
mPFC	Medial prefrontal cortex
MRI	Magnetic resonance imaging
NIMH	National Institute of Mental Health
PRT	Probabilistic Reward Task
QIDS-SR	Quick Inventory of Depression Scale – Self Report
SAE	Serious adverse event
SAP	Statistical analytical plan
SD	Standard deviation
SLIPS	Specific Loss of Interest and Pleasure Scale (SLIPS)
TEPS	Temporal Experience of Pleasure Scale
V0	Visit 0 (baseline)
V5	Visit 5 (primary outcome assessment at week 5 post-randomization)
VS	Ventral striatum
WHODAS	WHO Disability Assessment Schedule
WHO	World Health Organization

PURPOSE OF STATISTICAL ANALYTICAL PLAN (SAP)

The purpose of this SAP is to outline the planned analyses to be completed for the EZOR61-01 trial. The planned analyses identified in this SAP will be included in abstracts and manuscripts reporting the results of the trial. Exploratory analyses not necessarily identified in this SAP may also be performed. Any post hoc, or unplanned, analyses not explicitly identified in this SAP will be clearly identified as such in any published papers from this study.

1. INTRODUCTION

Depressive disorders are among the most disabling medical conditions worldwide and currently available treatments fall short of addressing this large public health burden. Dysfunction within the brain reward system is emerging as a core feature of depressive disorders, in particular related to deficits in motivation, interest, and response to pleasure (e.g., anhedonia: markedly diminished response to pleasure). Evidence from a series of preclinical studies highlighted the KCNQ subtype of neuronal potassium (K+) channel as a novel target for the treatment of depressive disorders, and data from human pilot study showed a reduction in anhedonia and related symptoms with an increased brain response to reward (as measured by functional magnetic resonance imaging [fMRI]) following treatment with ezogabine, a KCNQ channel opener. Building on these data, this study will assess reward circuit activity following treatment with ezogabine in depressed patients with anhedonia and will examine the relationship between change in reward circuit activity and clinically relevant symptom and behavior outcomes.

1.1 Study Objectives

1.1.1 Primary Objective

The primary objective of this study is to establish target engagement of a KCNQ channel opener (ezogabine) in patients with depressive disorder and anhedonia.

1.1.2 Secondary Objectives

Secondary objectives of this study are to determine if treatment with ezogabine is associated with a reduction in clinical symptoms of depression and anhedonia and with a modulation of behavioral responses to reward.

1.2 Study Design

This study is a Phase IIa, prospective, multi-site, randomized (1:1), parallel arm, double-blind, placebo-controlled clinical trial.

1.2.1 Study duration and time points

Participants will be followed for up to 14 weeks. The screening period may last up to 6 weeks; after randomization, titration and treatment lasts 5 weeks at which point the primary endpoint is ascertained. Participants will continue to be followed for a 3 week medication taper period after collection of the primary endpoint.

1.2.2 Randomization and masking

Patients will be randomized in a 1:1 allocation to ezogabine or placebo. Randomization will be stratified by center balanced by randomly permuted blocks, and treatment code will be allocated through the interactive web response system (IWRS). Randomization will take place only after it has been determined that all study eligibility criteria have been met and the baseline fMRI scan has been obtained.

This is a double-blind, placebo-controlled trial. Site personnel and patients will be blinded to treatment assignment throughout the trial.

2. STUDY POPULATIONS

Intent-to-Treat (ITT) Population
The ITT population is defined as all randomized subjects, regardless of when they withdrew from the study or received the assigned treatment. The ITT population will be used to present all data (including the primary endpoint) by randomized treatment group. Subjects will be analyzed according to the treatment to which they were randomized, regardless of the treatment they actually received.

3. EFFICACY ENDPOINTS

3.1 Primary Endpoint
The primary endpoint of this study is the baseline-corrected fMRI activation in an a priori bilateral ventral striatum (VS) area mask during the "gain" cue of the Incentive Flanker Task (IFT) contrasted with neutral cue for reward anticipation. The primary outcome fMRI will be obtained at 5 weeks post-randomization. The primary endpoint will be used as the sole criterion for the "go-no go" decision.

3.2 Secondary Endpoints
Secondary endpoints are measured at 5 weeks post-randomization unless otherwise noted. The following secondary endpoints will be analyzed. All secondary analyses will be considered exploratory and will not influence the study "go-no/go" decision.

3.2.1 fMRI Activation within the Reward Circuit
In the same manner as the primary endpoint, fMRI activation in other regions of the reward cortico-striatal circuit (medial prefrontal cortex [mPFC], bilateral dorsal striatum [DS], and anterior cingulate cortex [ACC]), as well as whole-brain analyses, will be measured during the gain cue of the IFT and contrasted with neutral cue.

Similarly, activation during the loss cues of the IFT task will be compared with the fMRI activation in neutral tasks.

3.2.2 Change in Anhedonia
Symptoms of anhedonia will be measured with the Snaith-Hamilton Pleasure Scale (SHAPS), the Temporal Experience of Pleasure Scale (TEPS), the Anticipatory and Consummatory Interpersonal Pleasure Scale (ACIPS), and the Specific Loss of Interest and Pleasure Scale (SLIPS). The SHAPS, TEPS, ACIPS, and SLIPS will be collected at baseline and weekly through the primary outcome assessment (V0 – V5).

3.2.3 Change in Depressive Symptoms
Symptoms of depression will be measured with the Quick Inventory of Depressive Symptomatology, Self-Report (QIDS-SR), the Montgomery-Asberg Depression Rating Scale (MADRS), the Mood and Anxiety Symptoms Questionnaire (MASQ),

and the Columbia-Suicide Severity Rating Scale (C-SSRS). The QIDS, MADRS, and C-SSRS will be collected at baseline (V0) and weekly through the primary outcome assessment (V1 – V5). The MASQ will be collected at baseline (V0) and primary outcome assessment (V5) only.

3.2.4 General Functioning

The Clinical Global Impression – Improvement and Clinical Global Impression – Severity scales will be used to measure overall illness severity at baseline and weekly through the primary outcome assessment (V0 – V5).

The World Health Organization (WHO) Disability Assessment Schedule 2.0 (WHODAS 2.0) will also measure general functioning. The WHODAS will be measured at baseline (V0) and at the time of the primary outcome assessment (V5).

3.2.5 Reward learning

The Probabilistic Reward Task (PRT) will be used as a behavioral assessment to calculate an index of reward learning at baseline (V0) and at the time of the primary outcome assessment (V5).

4. SAFETY ENDPOINTS

The incidence and frequency of all anticipated and unanticipated serious and non-serious adverse events (AEs) that occur between randomization and study exit (V8) eight weeks after randomization will be determined. Severity/grade and expectedness will be assessed by the study clinician. The Medical Dictionary for Regulatory Activities (MedDRA) system organ class and preferred terms will be recorded for each reported AE.

5. STATISTICAL METHODOLOGY

5.1 Statistical Methods

All data will be presented by treatment group as well as for the whole sample.

For all baseline, demographic and efficacy variables, data will be summarized and analyzed by randomized treatment group as per ITT.

In summary tables of normally distributed continuous variables, the following descriptive statistics will be used: minimum and maximum, mean, and standard deviation (SD). In summary tables of non-normally distributed variables data will be presented using the minimum, maximum, median and lower and upper quartiles.

In summary tables of categorical variables, counts and percentages will be used. The denominator for each percentage will be the number of patients within the randomization group unless otherwise specified.

Rates of events will be calculated as the ratio of the total number of events recorded divided by the total patient-time in study (months or days). Total patient-time will be

calculated by summing the time (in days) that patients were at risk for a specific event from the time they were randomized in the study until study exit. Rates and 95% confidence intervals will be reported.

For any variable measured at multiple points in time, change from baseline will be calculated as the difference between the value of the variable at a specific point in time (e.g. 5 weeks) minus the baseline value.

Relative change from baseline will be calculated as the value of a parameter at a specific point in time minus the baseline value of the parameter divided by the baseline value of the parameter.

Percent change will be calculated as the relative change multiplied by 100.

Hypothesis testing will be conducted at the 0.05 two-sided significance level for the primary endpoint. No formal hypothesis tests will be conducted for the secondary endpoints, rather, we will calculate estimates and 95% confidence intervals by group.

P-values will be rounded to three decimal places. P-values less than 0.001 will be reported as <0.001 in tables. P-values greater than 0.999 will be reported as >0.999.

Should any of the statistical methods proposed prove unsuitable during data analysis, more appropriate methods will be used. These include data transformation (for example to a logarithmic scale) to satisfy model assumptions such as normally distributed residuals with constant variance, or the application of non-parametric techniques.

Additional ad-hoc analyses may be conducted as deemed appropriate.

All statistical analysis will be performed using SAS V9.4 or higher.

5.2 Handling of Dropouts and Missing Data

Missing baseline values that are needed to compute absolute, relative, or percent change from baseline will be imputed using mean imputation (Thompson, 2005). The missing values of a variable will be replaced with the observed sample mean of that variable. Mean imputation is appropriate because baseline variables are independent of randomization assignment.

5.2.1 Missing values for the primary outcome

Missing outcome values for primary endpoint analysis will not be imputed. For the primary endpoint, missing data can be the result of patient refusal, missed visit, or study dropout (withdrawal, lost to follow-up). Missing data due to patient refusal or missed visits are assumed to be missing at random (MAR). Study dropout is also assumed to be MAR.

5.2.2 Missing values for the secondary outcomes

Missing outcome values for secondary endpoints will not be imputed.

5.2.3 Sensitivity analysis

Sensitivity analyses will be conducted to investigate possible violations of the missing at random assumption. Specifically, we will determine if patients missing the primary outcome fMRI differ from patients with fMRI data at V5 in terms of:

- Baseline demographics
- Baseline rewards circuit activation
- Longitudinal anhedonia symptoms
- Longitudinal depression symptoms
- Treatment group

Three imputed datasets will be created to assess the robustness of the primary endpoint analysis:

- Multiple imputation will be used to impute any missing V5 value of the bilateral VS activation
- Lowest overall observed (across both treatment groups) of the bilateral VS activation will be imputed for any missing V5 value
- Highest overall observed (across both treatment groups) of the bilateral VS activation will be imputed for any missing V5 value

In addition, we will conduct sensitivity analyses to examine the effect of lowering the dose of ezogabine in patients who could not tolerate the protocol-defined maximum dosage. Specifically, we will determine if patients on a reduced dose compared to the protocol-defined dose differed in terms of

- Baseline demographics
- Baseline rewards circuit activation
- Longitudinal anhedonia symptoms
- Longitudinal depression symptoms

We will repeat the primary endpoint analyses in the subset of patients treated per-protocol and compare primary endpoint results between patients on a reduced dose versus the protocol-defined maximum dose.

5.3 Patient Characteristics

5.3.1 Patient Disposition

The subject disposition table will summarize patients' characteristics by randomization group and overall.

- The number (%) of patients randomized
- The number (%) of patients withdrawn or lost to follow-up by the primary outcome visit (week 5)

The number (%) of subjects who complete and withdraw/lost to follow-up from the study and the primary reason for withdrawal will be summarized by randomization group and

overall for all subjects.

The percentages will be calculated based on the total number of patients randomized in each randomization group.

5.3.2 Protocol Deviations

Major and minor protocol deviations are defined as deviations from the procedures outlined in the protocol. All statistical analyses and summaries will be conducted on an intent-to-treat basis (with the exception of the sensitivity analysis to examine the effect of dose reduction due to medication intolerance).

Treatment compliance will be compared by treatment group with compliance defined as the number (%) of total prescribed pills taken, and subject report will be compared to serum drug levels.

5.3.3 Demographic Characteristics

Demographic and baseline characteristics data including age, gender, race, and ethnicity will be summarized using summary statistics for continuous variables or by group frequencies and percentages for categorical variables, as appropriate.

5.4 Efficacy Analysis

5.4.1 Analysis of Primary Endpoint and Determination of Sample Size

The primary endpoint is the change in activation of the bilateral VS, a region of the rewards circuit in the brain, at 5 weeks from randomization. Activation will be measured by fMRI, and beta weights for each subject will be calculated to create a contrast between gain and neutral cues to model reward anticipation as described in the Neuroimaging Methods section of the protocol. To calculate the beta weight for the bilateral VS, the beta weights for the right and left VS will be averaged.

A linear mixed effects model with a single random intercept term will be used to assess the change in activation of the bilateral VS by treatment group. A sample size of 48 provides 85% power to detect a difference in means between groups of 0.8 standard deviations assuming a correlation of 0.6 between a pair of measurements made on the same subject using a two-sided 0.05 significance level test.

Although randomization is stratified by site, the primary endpoint analysis will be of the pooled dataset; secondary analyses will provide site-specific estimates.

No adjustment to the primary analysis will be made for baseline measures that are found to be imbalanced but, if applicable, secondary sensitivity analyses may adjust for any factors that are imbalanced at a significance level below $p = 0.05$. (Altman, 1985) (Peter Peduzzi, 2002) (Senn, 1989) (Senn, Testing for baseline balance in clinical trials, 1993)

5.4.2 Secondary Endpoints

All secondary analyses will be performed using the ITT population.

As the secondary fMRI outcomes are exploratory analyses, we will present the mean and 95% confidence interval of the difference in change in activation between treatment groups rather than conducting formal hypothesis testing. As with the bilateral VS, a single beta weight for the bilateral DS will be calculated by averaging the beta weights for the left DS and right DS.

Linear mixed effects models will be used to analyze the anhedonia, depression, and general functioning scales that are collected at multiple time points. Slopes will be estimated with 95% confidence intervals by treatment group. Mean change and 95% confidence intervals by group will be presented for measures collected at baseline and primary outcome only.

5.5 Safety Analysis

All adverse events in this study will be analyzed from the time of randomization. All adverse events will be analyzed using the ITT population.

AEs will be mapped to the MedDRA coding system. The incidence of non-serious AEs and serious AEs will be summarized by System Organ Class, MedDRA Preferred Term and treatment**.**

The number and percentage of patients with an AE or serious AE will be presented. For the incidence count, a patient will be counted only once for a specific MedDRA Preferred Term, although a MedDRA Preferred Term might be recorded more than once for a particular patient.

The number and rate of each AE will also be presented. Differences in the incidence of individual AEs will be assessed using negative binomial regression with the treatment group as the predictor.

5.6 Interim Analysis

There is no planned interim analysis for this study.

6. REFERENCES

Altman, D. G. (1985). Comparability of randomized groups. *Statistician*, 125-136.

Peter Peduzzi, W. H. (2002). Analysis of Randomized Controlled Trials. *Epidemiologic Reviews*, 26-38.

Senn, S. (1989). Covariate imbalance and random allocation in clinical trials. *Statistics in Medicine*, 467-475.

Senn, S. (1993). Testing for baseline balance in clinical trials. *Statistics in Medicine*, 1715 - 1726.

Thompson, I. R. (2005). Adjusting for partially missing baseline measurements in randomized trials. *Statistics in Medicine*, 993-1007.