Title: An Open-Label, Single-Arm, Multicenter, Prospective, Phase 4, Interventional, Flexible Dose Study to Evaluate the Effectiveness of Vortioxetine on Goal Achievement After a Change in Antidepressant Medication for the Treatment of Subjects With Major Depressive Disorder

NCT Number: NCT02972632

SAP Approve Date: 15 February 2018

Certain information within this Statistical Analysis Plan has been redacted (ie, specific content is masked irreversibly from view with a black/blue bar) to protect either personally identifiable information or company confidential information.

This may include, but is not limited to, redaction of the following:

- Named persons or organizations associated with the study.
- Proprietary information, such as scales or coding systems, which are considered confidential information under prior agreements with license holder.
- Other information as needed to protect confidentiality of Takeda or partners, personal information, or to otherwise protect the integrity of the clinical study.

STATISTICAL ANALYSIS PLAN

STUDY NUMBER: Vortioxetine-4003

An Open-Label, Single-Arm, Multicenter, Prospective, Phase 4, Interventional, Flexible Dose Study to Evaluate the Effectiveness of Vortioxetine on Goal Achievement After a Change in Antidepressant Medication for the Treatment of Subjects With Major Depressive Disorder

Goal Achievement After a Change to Vortioxetine in Adults With Major Depressive Disorder

PHASE 4

Version**:** 2.0

Date**:** 15 February 2018

Prepared by:

Personal Protected Data

Based on:

Protocol Version: 1.0

Protocol Date: 18 August 2017

1.1 Approval Signatures

Study Title: **An Open-Label, Single-Arm, Multicenter, Prospective, Phase 4, Interventional, Flexible Dose Study to Evaluate the Effectiveness of Vortioxetine on Goal Achievement After a Change in Antidepressant Medication for the Treatment of Subjects With Major Depressive Disorder**

Approvals:

Personal Protected Data

26 Feb. 2018
Date

26 Feb 2018
Date

26 Feb 2018
Date

26 Feb 2018
Date

27 Feb 2018
Date

2.0 TABLE OF CONTENTS

LIST OF IN-TEXT TABLES

Appendix A. Criteria for Identification of Markedly Abnormal Values for Vital Signs

LIST OF IN-TEXT FIGURES

Figure 2.a Schematic of Study Design

3.0 LIST OF ABBREVIATIONS

5-HT	Serotonin
AE	adverse event
ALT	alanine aminotransferase
AST	aspartate aminotransferase
CGI-I	Clinical Global Impression Scale-Improvement
CGI-S	Clinical Global Impression Scale Severity
CRO	contract research organization
C-SSRS	Columbia Suicide Severity Rating Scale
DRESS	drug reaction with eosinophilia and systemic symptoms
DSM-5	Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
Company Confidential Information	
ECG	Electrocardiogram
eCRF	electronic case report form
FAS	full analysis set
FSH	follicle-stimulating hormone
GAS	Goal Attainment Scale
GCP	Good Clinical Practice
hCG	human chorionic gonadotropin
ICH	International Conference on Harmonisation
ID	Identification
IEC	independent ethics committee
INR	international normalized ratio
IRB	institutional review board
IxRT	Interactive Response Technology
Company Confidential Information	
LFT	liver function tests
MDD	major depressive disorder
MDE	major depressive episode
MedDRA	Medical Dictionary for Regulatory Activities
PGx	Pharmacogenomics
PDQ-D	Perceived Deficits Questionnaire-Depression
PHQ-9	Patient Health Questionnaire
PPS	per protocol set
PTE	pretreatment event
Q-LES-Q	Quality of Life and Enjoyment and Satisfaction Questionnaire
QD	once daily
Company Confidential Information	
SAE	serious adverse event
SJS	Stevens-Johnson syndrome
SUSAR	suspected unexpected serious adverse reaction

TEN	toxic epidermal necrolysis
UDS	urine drug screen
ULN	upper limit of normal
WHO	World Health Organization
WHO-5	5-item World Health Organization Well-Being Index

4.0 OBJECTIVES

4.1 Primary Objective

To determine the effectiveness of treatment with vortioxetine on subject goal achievement after a change in antidepressant medication for the treatment of major depressive disorder (MDD).

4.2 Secondary Objectives

To determine the effectiveness of treatment with vortioxetine on depressive symptoms/outcomes, clinical global impression, cognitive impairment, and quality of life in subjects with MDD.

4.3 Additional Objectives

- To determine the effectiveness of treatment with vortioxetine on cognitive performance and functionality in subjects with MDD.
- To determine the effectiveness of treatment with vortioxetine on work performance in subjects with MDD.
- To determine the effectiveness of treatment with vortioxetine on general health status and well-being in subjects with MDD.
- To assess health care utilization and the safety and tolerability of vortioxetine in subjects with MDD.

4.4 Study Design

This study is an open-label, single-arm, multicenter, prospective, phase 4, interventional, flexible dose study (as per the Trintellix US package insert) to determine the effectiveness of treatment with vortioxetine on subject goal achievement after a change in antidepressant medication for the treatment of MDD.

This study is comprised of a 3-week Screening Period (maximum 21 days), to be followed by a 12-week Treatment Period and a 4-week Safety Follow-up Period. Study visits will occur at Baseline (Day 1), and Weeks 2, 4 (Telephone), 6, 9 (Telephone), and 12. This study will enroll approximately 120 subjects and will be conducted at approximately 25 sites in the United States.

This study will enroll men and women between the ages of 18 and 65, inclusive, with a primary diagnosis of MDD, who have previously been treated within 6 weeks of Screening or are currently being treated for a major depressive episode (MDE) with an approved antidepressant for 6 weeks or longer at an adequate dose. Subjects who are dissatisfied with their current or recent antidepressant and are interested in a change in antidepressant treatment will be eligible for Screening.

The change in antidepressant medication will be based on the investigator's judgment in collaboration with the subject's level of dissatisfaction with their current or recent antidepressant treatment (ie, residual mood symptoms, intolerable side effects). Subjects also need to be eligible

per the Trintellix US package insert and sign and date a written, informed consent form and any required privacy authorization prior to the initiation of any study procedures.

The investigator or the investigator's designee will access the Interactive Response Technology (IxRT) at Screening to obtain the subject study number. At the Baseline Visit, if the subject has satisfied all of the inclusion criteria and none of the exclusion criteria for enrollment, the subject will be enrolled using the IxRT. Subjects will be instructed on when to take the first dose of study drug. The dosing (both initial and adjustments) will be determined by investigator or designee as per routine clinical practice and per approved Trintellix labeling. Subjects will determine treatment goals with their clinicians using the Goal Attainment Scale (GAS) adapted for an MDD population. The subject and clinician will work to establish 3 goals at Baseline. One goal will be determined based on the subject's self-defined objectives. Two goals will be selected from the following predefined categories representing MDD residual symptoms, antidepressant side effects, and common reasons for a change in antidepressant medication, including but not limited to:

MDD Domains	Psychological	Motivation	Emotional	Physical/ Functional	Cognition
Predefined goal categories:	-Depressed mood -Anxious mood	-Lack of motivation	-Low self esteem -Anhedonia	-Fatigue -Insomnia -Muscle tension -Sexual dysfunction -Weight gain	-Problems concentrating -Short term memory problems -Difficulty staying focused -Difficulty paying attention -Problems thinking clearly

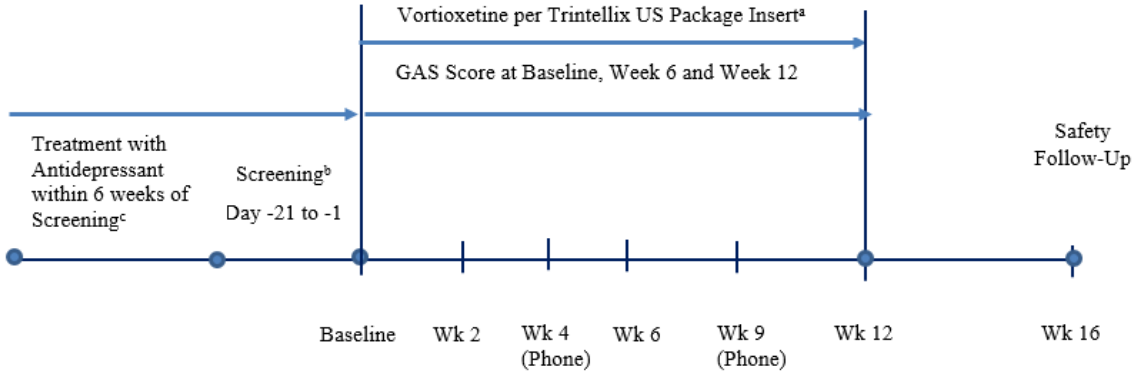
A prepopulated list of categories will be provided in the structured guidance for goal assessment, including behavioral examples describing goals and observable benchmarks to help clinicians set the appropriate goals with subjects. Goals will be set at Baseline/Visit 1. Progress will be evaluated at the Week 6 visit. At the Final Visit the goals will be assessed based on levels of achievement on a 5-point scale. Investigators will be trained on goal setting/goal assessment and monitored for ability in key domains of goal setting (eg, attainability of goals, use of observable objectives and benchmarks, equidistance of scaling, general level of difficulty, goal differentiation, and overall quality).

A safety follow-up phone call will be made 4 weeks after completion of the 12 weeks of treatment. Subjects who discontinue or who are withdrawn prior to study completion will come to the site for an Early Termination Visit as soon as possible and will be contacted for a safety follow-up 4 weeks after the last dose of study medication.

Enrollment caps may be implemented throughout the course of the study based on the patient reported outcome (ie, Patient Health Questionnaire (PHQ-9) and Perceived Deficits Questionnaire-Depression (PDQ-D)) scores at Baseline.

A futility analysis will be performed at an interim analysis with approximately 60 subjects completing Week 12. If the futility analysis indicates that the study should be discontinued, the final analyses will be conducted for all enrolled subjects. Details for the futility analyses are in the Interim Analysis section of this SAP.

Figure 4.a Schematic of Study Design



(a) Recommended starting dosage is 10 mg/day as ideal dosage. Dosage may be decreased to 5 mg/day at discretion of the investigator or designee.
(b) If all screening activities are completed and the subject meets the eligibility criteria the Screening and Baseline Visits can be combined into 1 visit.
(c) Subjects currently on an antidepressant at Screening will be discontinued in a manner that is consistent with labeling recommendations and conventional medical practice prior to Baseline.

5.0 ANALYSIS ENDPOINTS

The GAS, adapted for MDD is being used as primary endpoint in this study. As this is an adaptation of an existing instrument the results will provide novel outcome data for the goal attainment literature. Therefore, secondary endpoint assessments with validated instruments will provide additional confirmation of the effectiveness of vortioxetine in treating MDD when current antidepressant treatment is not adequate.

The following outcome measures will be used for secondary and exploratory endpoints as they are frequently used to assess clinical response: PHQ-9, PDQ-D, Quality of Life and Enjoyment and Satisfaction Questionnaire (Q-LES-Q), 5-item World Health Organization Well-Being Index (WHO-5), Clinical Global Impression-Severity (CGI-S), Clinical Global Impression-Improvement (CGI-I), Company Confidential Information [redacted]

Measures have been taken regarding the methodology of this study to assess suicidal risk. The selection criteria exclude the participation of subjects at significant risk for suicide. Throughout the study, signs of suicidal risk will be assessed both by rating scale assessment (C-SSRS) and by investigators' clinical judgment. Subjects will be withdrawn from the study in case of such risk. Furthermore, subjects will be screened for the history of suicidal behavior.

5.1 Primary Endpoint

The primary endpoint is to estimate the proportion of subjects who achieve their identified goals as demonstrated by an outcome Goal Attainment Scale (GAS) Score of $\geq$50 at Week 12.

5.2 Secondary Endpoints

The secondary endpoints are:

- Mean change from Baseline in total GAS Score at Week 6 and Week 12.
- Change from Baseline in depression severity as measured by the Patient Health Questionnaire (PHQ-9) Score at Week 6 and 12.
- Change from Baseline in perceived cognitive deficits and cognitive function as measured by the Perceived Deficits Questionnaire-Depression (PDQ-D) at Week 6 and 12.
- Change from Baseline in Quality of Life Enjoyment and Satisfaction Scale (Q-LES-Q) Score at Week 12.
- Change from Baseline in 5-item World Health Organization Well-Being Index (WHO-5) Score at Week 12.
- Change from Baseline in Clinical Global Impression Scale Severity (CGI-S) at Week 12.
- Clinical Global Impression Scale-Improvement (CGI-I) Score at Week 12.

5.3 Additional Endpoints

The exploratory/additional endpoints include:

5.4 Safety Assessments

Safety and tolerability of vortioxetine will also be evaluated using the following general assessments:

- Adverse events (AEs).
- AEs leading to discontinuation.
- Weight.
- Columbia-Suicide Severity Rating Scale (C-SSRS).

6.0 DETERMINATION OF SAMPLE SIZE

As no formal hypothesis testing is performed for the primary endpoint, the purpose of sample size estimation is to determine the adequate sample size needed to provide a precise estimate for the response rate (GAS Score ≥50) with 95% CIs.

Conservatively estimating that 50% of subjects will be responders (GAS≥50) and that 100 subjects will complete the study, CIs will be ±0.098 around the point estimate (proportion of responders).

7.0 METHODS OF ANALYSIS AND PRESENTATION

7.1 General Principles

This is an open-label study with a single treatment regimen and all comparisons are made to the conditions prior to the subject being switched to vortioxetine.

Baseline values are defined as the last observed value before the first dose of study switch medication.

All statistical analyses will be conducted using SAS® Version 9.2, or higher.

All confidence intervals, statistical tests, and resulting P-values will be reported as 2-sided and will be assessed at α=0.05 significance level unless otherwise stated. P-values will be rounded to 3 decimal places prior to assessment of statistical significance.

Means and medians will be presented to 1 more decimal place than the recorded data. The standard deviations (SDs) will be presented to 2 more decimal places than the recorded data. Confidence intervals about a parameter estimate will be presented using the same number of decimal places as the parameter estimate.

Where appropriate, variables will be summarized descriptively by study visit. For the categorical variables, the count and proportions of each possible value will be tabulated. The denominator for the proportion will be based on the number of subjects who provided non-missing responses for the categorical variable. For continuous variables, the numbers of subjects with non-missing values, mean, median, SD, minimum, and maximum values will be tabulated.

7.1.1 Definition of Study Days

Study Day 1 is defined as the date on which a subject is administered their first dose of the vortioxetine medication. Other study days are defined relative to the Study Day 1 with Day 1 being Study Day 1 and Day -1 being the day prior to Study Day 1.

7.1.2 Definition of Visit Window

The window definitions as stated will be applicable for all the summaries where study visit data is presented and where study visit data are compared with baseline data. The window conventions for safety and efficacy parameters are consecutive; that is, the lower or upper bound of the windows is the midpoint between 2 consecutive study visits.

One or more results for a parameter may be obtained in a visit window. In such an event, the result with the closest date to the scheduled visit date will be used. In the event of 2 observations within a given visit window being equidistant to the scheduled visit date, the later observation will be used.

7.2 Analysis Sets

In this study, 3 kinds of analysis sets are defined: full analysis set (FAS), per protocol set (PPS) and safety analysis set.

The FAS, used for primary effectiveness analysis, will include all subjects who were enrolled in the treatment period and received at least 1 dose of the study drug and completed at least 1 GAS assessment post-Baseline. In FAS effectiveness summaries, subjects will be considered as vortioxetine-treated no matter what dose levels were received during the study.

A PPS will include all FAS subjects who had no major protocol violations. Subjects to be excluded from the PPS, whether due to protocol violations or noncompliance to the dosing prescribed by the treating physician, will be identified in the minutes of the subject evaluability assessment.

The Safety Analysis Set will include all subjects who were enrolled and received at least 1 dose of study medication. In safety summaries, all subjects will be analyzed as vortioxetine-treated.

The sponsor will verify the validity of the definitions of the analysis sets as well as the rules for handling data, with consulting a medical expert as needed. If necessary, the Handling Rules for Analysis Data will be supplemented with new handling rules that were not discussed at the planning stage. The Handling Rules for Analysis Data must be finalized prior to database lock.

7.3 Disposition of Subjects

A subject disposition summary will be provided. The categories will include all subjects who were enrolled, subjects who were not treated, subjects who discontinued from the study categorized by reason, and subjects who completed the study. The primary reasons for discontinuation will include AE, major protocol deviation, lost to follow-up, study termination, pregnancy, lack of efficacy, voluntary withdrawal, and other.

A summary of screening failures and listings of inclusion/exclusion criteria responses for subjects with violations will also be provided.

7.4 Demographic and Other Baseline Characteristics

Baseline characteristics will be listed and summarized for demographics (sex, age, race, and body mass index), and medical history for the safety analysis set.

Categorical variables will be summarized using frequencies and percentages. Continuous variables will be summarized using mean, SD, median, maximum and minimum values.

7.5 Medical History and Concurrent Medical Conditions

Medical history refers to the significant conditions/diseases that stopped at or prior to signing of informed consent. Concurrent medical conditions are those significant ongoing conditions/diseases present at signing of informed consent.

Medical history and psychiatric history and concurrent medical conditions will be coded using the latest version of Medical Dictionary for Regulatory Activities (MedDRA) and will be

summarized in tables using System Organ Class (SOC) and MedDRA preferred term. The tables will include numbers and percentages of subjects and will be sorted in alphabetical order by system organ class and preferred term. A subject will only be counted once within a particular class even if he/she has multiple conditions/symptoms. Summaries will be based on the safety analysis set.

All medical history and concurrent medical condition data will be presented in listings.

7.6 Medication History and Concomitant Medications

The medication history and concomitant medications are defined as follows:

- Medication history refers to the medication that the study subjects stopped taking within 30 days prior to signing of informed consent (ie, stop date < informed consent date).
- Concomitant medication is defined as medication that the study subjects continued taking or took from signing of informed consent through end of study:
  - Concomitant medication that started and stopped prior to baseline (ie, stop date ≥ informed consent date, and stop date ≤ first dose date).
  - Concomitant medication that started prior to and was ongoing at baseline (ie, start date < first dose date, and stop date > first dose date) or if coded as ongoing with no stop date.
  - Concomitant medication that started after baseline but before or at last dose (ie, start date ≥ first dose date, and start date ≤ last dose date).
  - Concomitant medication that started prior to and was ongoing at baseline or that started after baseline (ie, start date ≤ last dose date, and stop date > first dose date.
  - If start date and stop date are missing, medication will be assumed to occur both prior and concomitantly.

Medication history and concomitant medications will be coded using the latest version of the World Health Organization (WHO) Drug Dictionary and summarized by giving the number and percentage of subjects by preferred term within each therapeutic class, with therapeutic class and medications in each class sorted in alphabetical order. The total number of subjects with medications in each selected therapeutic class will also be presented. If a subject reports taking 2 drugs belonging to the same class, he/she will only be counted once within that class. Summaries of medication history and concomitant medication will be based on the safety analysis set.

All prior and concomitant medications data will be presented in listings.

7.7 Study Drug Exposure and Compliance

The summary of study drug exposure and compliance will be based on the safety analysis set.

Duration of exposure to study medication is defined as (date of last dose – date of first dose +1). Treatment duration will be summarized by duration category in weeks (1 to 2 weeks, 3 to 4 weeks, 5 to 6 weeks, 7 to 8 weeks, 9 to 10 weeks, and 11 to >12 weeks) and the number and

percentage of subjects in each duration category. Treatment duration (days) will also be summarized using descriptive statistics (n, mean, SD, median, minimum, and maximum).

Percent of study drug compliance is defined as {(number of tablets dispensed – number of tablets returned)/(date of last dose – date of first dose +1)}x100%. If a value for the number of returned tablets is missing or the return date is missing, then 100% compliance will be assigned for each day up to the number of tablets dispensed or up to the date of return whichever is earlier.

Study medication compliance will be summarized by compliance category (<80%, 80 to 120%, and >120%) and the number of subjects in each compliance category. Study medication compliance will also be summarized as a continuous variable using descriptive statistics (n, mean, SD, median, minimum, and maximum).

All study drug administration and accountability data will be listed by study site and subject number. The following variables will be listed: subject identifier, visit number, first and last dose dates, medication identification number, date dispensed and returned, number of tablets dispensed and returned, and percent compliance.

7.8 Efficacy Analysis

The primary objective of the study is to determine the effectiveness of treatment with vortioxetine on subject goal achievement after changing antidepressant medication for treatment of MDD. The primary endpoint is the proportion of subjects who changed to treatment with vortioxetine and achieved their individually identified goals as demonstrated by an outcome GAS Score of ≥50 at Week 12. This proportion will be summarized at Week 12. All efficacy analyses will use the FAS.

7.8.1 Primary Efficacy Endpoint

The GAS is a tool to measure the progress each subject has towards achieving their individualized goals. While the outcome measure is unique to each subject, standardized scoring is applied to allow for statistical analysis.

The GAS allows for 1 of 2 types of goals to be set:

(1) Subject defined goals allow a "subject's chief complaint" approach taken towards establishing the goal, with no specific limitations set for the type or focus of the goal, beyond requiring that it meets the basic standards of measurability, equidistance, and difficulty.

(2) Domain-defined goals are more structured in nature, presenting the subject with an array of broadly defined domains and more specific subdomains. After selecting an appropriate domain and (where applicable) subdomain, the clinician will work with the subject to establish a specific goal that fits within the selected category.

For GAS, a total of 3 goals will be selected, 1 subject defined and 2 from pre-defined domain categories. A 5-point rating scale will then be applied to each of the 3 goals. A semi-structured interview will be conducted with each subject to conduct goal-setting at the outset of the study.

Another evaluation will take place at Week 6 and Week 12/End of Study visits or at Early Termination Visit (as applicable) to determine the level of progress at that time, with assignment of appropriate scores. It takes subjects approximately 45 minutes to complete the GAS.

Calculating the Overall GAS

The goals identified are incorporated into the single GAS score by applying a formula:

$$\text{Overall GAS} = 50 + \frac{10*\sum(w_i x_i)}{\sqrt{(1-\rho)\sum w_i^2 + \rho(\sum w_i)^2}}$$

Where:
w_i = the weight assigned to the *i*th goal (if equal weights, $w_i = 1$)
x_i = the numerical value achieved (between –2 and + 2)
ρ = the expected correlation of the goal scales
(For practical purposes, most commonly approximates to ρ=0.3)

For this study, a choice of weight is based on importance of the goal. Weight is 1, 2, and 3 for least important, moderately important and most important goals, respectively. The expected correlation of 0.3 was selected for this study.

The GAS will be summarized for both Week 6 and Week 12 using both percents of response and as a continuous score. An overall GAS score greater than 50 is considered as a response.

7.8.2 Secondary Efficacy Endpoint(s)

7.8.2.1 PHQ-9

The PHQ-9 will be used as a secondary efficacy measure of depression severity. This is a well-established patient reported outcome tool for assessment of change in depressive symptoms and is a sensitive measure of depression severity.

The PHQ-9 consists of a 9-item scale originally developed for primary care settings, with 1 item corresponding to each of the 9 MDE symptom criteria for depression in DSM, asking if they have bothered the patient over the last 2 weeks. Each question is rated on a scale from 0 (not at all) to 3 (nearly every day). If any problems are answered 1 or higher, a final question on how difficult those problems made it to do work, take care of things at home, or get along with other people is completed rated from not difficult at all to extremely difficult. The 9 questions are summed to a total score ranging from 0 to 27 with higher scores reflecting greater severity. It

takes approximately 5 minutes to complete the PHQ-9. The total score for the PHQ-9 is analyzed as a continuous variable.

7.8.2.2 PDQ-D

The PDQ-D, a validated version of the original PDQ specifically modified to depressed patients, will be used to evaluate self-reported cognitive complaints in patients with MDD. The PDQ was originally developed to assess cognitive complaints in patients with multiple sclerosis. However, it does address cognitive aspects highly relevant for depression, including attention and concentration, retrospective memory, prospective memory, and planning and organization. It takes approximately 10 minutes to complete the PDQ-D. The score from the PDQ-D is analyzed as a continuous variable.

7.8.2.3 Clinical Global Impression Scales

The CGI was developed to provide global measures of the severity of a patient's clinical condition during clinical studies. The CGI scales consist of 2 subscales: the CGI-S and the CGI-I. CGI-S and CGI-I raters will be required to have suitable experience and training (as deemed by sponsor) to be eligible to rate subjects within the study. The total time to score the CGI is approximately 1 to 2 minutes after the clinical interview.

CGI-S

The CGI-S is a clinician rated scale designed to assess global severity of illness and change in the clinical condition over time. The CGI-S provides the clinician's impression of the patient's current state of mental illness. The clinician uses his or her clinical experience of this patient population to rate the severity of the patient's current mental illness on a 7-point scale ranging from 1 (normal - not at all ill) to 7 (among the most extremely ill patients). The CGI-S is analyzed as a continuous variable.

CGI-I

The CGI-I assesses the subject's improvement (or worsening). The clinician is required to assess the subject's condition relative to Baseline (Randomization/Day 1) on a 7-point scale ranging from 1 (much improved) to 7 (much worsened). In all cases, the assessment should be made independent of whether the rater believes the improvement/worsening is drug-related or not. The CGI-I is analyzed as a continuous variable.

7.8.2.4 Q-LES-Q

The Q-LES-Q is a scale designed to allow researchers to assess the degree of a subject's quality of life in various areas of daily living. Ninety-one of the 93 items are assembled into 8 categories: physical health/activities, feelings, work, household duties, school/course work, leisure time activities, social relations, and general activities. The items are grouped according to these categories and the items are each scored on a 5-point scale, from 1 (not at all or never) to 5 (frequently or all the time), to indicate the degree of enjoyment or satisfaction experienced.

Higher scores indicate greater enjoyment/satisfaction. Typically a self-report measure, this scale will be read to the subject by a member of the site staff and verbatim answers recorded to ensure all data points are captured. It takes approximately 40 to 45 minutes to complete the Q-LES-Q. Each of the 8 sub-scores from the Q-LES-Q is analyzed as a continuous variable.

7.8.2.5 WHO-5

The WHO-5 is a short, self-administered questionnaire covering 5 positively worded items, related to positive mood (good spirits, relaxation), vitality (being active and waking up fresh and rested), and general interests (being interested in things). Administering the WHO-5 takes 2 to 3 minutes to complete. The raw score is calculated by totaling the figures of the five answers. The raw score ranges from 0 to 25, 0 representing worst possible and 25 representing best possible quality of life. To obtain a percentage score ranging from 0 to 100, the raw score is multiplied by 4. A percentage score of 0 represents worst possible, whereas a score of 100 represents best possible quality of life.

7.8.3 Additional Efficacy Endpoint(s)

7.8.3.1 Compa ny

Company Confidential Information

[redacted]

7.8.3.2 Company Confidential

Company Confidential Information

[redacted]

7.8.3.3 Company

Company Confidential Information

[redacted]

Company Confidential Information

7.8.3.4 Company

Company Confidential Information

7.8.4 Analyses for Secondary and Additional Endpoints

For the secondary endpoints of outcome measures of GAS Score, PHQ-9, PDQ-D, Q-LES-Q Score, WHO-5, and CGI-S, total scores and changes from Baseline will be summarized by each time point of Baseline, Week 6 (as appropriate) and by each time point of Baseline and Week 12. These changes from Baseline will be analyzed using paired t-tests. Also, for each of the outcome measurements, change from Baseline at Week 12 will be analyzed using paired t-tests. Tests of significance will be 2-tailed with alpha level at 0.05. CGI-I Scores will be summarized at Week 12 using descriptive statistics.

Company Confidential Information

Additionally, data will be summarized by the following response and remission rates, respectively:

- response defined on the PHQ-9 (>=50% reduction in total score from Baseline) at week 6 and week 12
- response based on CGI-I (CGI-I score of <=2) at week 6 and week 12
- remission defined on PHQ-9 (PHQ-9 score of <=4) at week 6 and week 12
- remission based on CCI-S (CGI-S score of <=2) at week 6 and week 12.

Company Confidential Information

In order to assess Goal Attainment Scaling as a sensitive, reliable measure of response to drug treatment for depression, a series of analyses will be performed to examine the GAS approach against other measures assessed in the study, as well as to evaluate change in GAS scores over time. A full description of the analysis will be described in a separate validation plan.

7.9 Safety Analysis

7.9.1 Adverse Events

All adverse events will be coded using the latest version of MedDRA. In this dictionary, each verbatim term is coded to a lower level term, and then mapped to a preferred MedDRA term, which is then mapped to an SOC. All adverse events from the safety analysis set will be included in the data listings but only treatment-emergent adverse events will be included in the summary tables.

A treatment-emergent adverse event (TEAE) is defined as an adverse event with an onset that occurs after receiving study drug (AE start date ≥ first dose date) and within 30 days after receiving the last dose of study drug (AE start date – last dose date ≤30). Adverse events with onset occurring more than 30 days after last dose of study drug (AE start date – last dose date >30) will be listed, but not included in the summary tables. Adverse events with missing onset dates will be summarized regardless of severity and relationship to study medication.

Serious adverse events (SAEs) with onset that occurs after receiving study drug (AE start date ≥first dose date) and within 30 days after receiving the last dose of study drug (AE start date - last dose date ≤30) will be summarized.

In the high-level adverse event summary tables, TEAEs will be summarized regardless of intensity and relationship to study drug. Within each subject, multiple reports of events that map to a common MedDRA term will be counted only once for calculating incidence rates.

At the adverse event level, the summary tables will present the number of subjects reporting each of these MedDRA events, ie, the numbers of subjects reporting 1 or more events that map to the given MedDRA term.

At the SOC level, the summary tables will present the number of subjects reporting 1 or more events that map to the given SOC. That is, the number of subjects reported at the SOC level will be less than or equal to the sum of the subject counts across all adverse events within that SOC.

In selected summaries (TEAEs overview, and TEAEs by SOC and preferred term), adverse events will be summarized by the number of events reported in addition to the number and percentage of subjects with at least one event.

For the summary of TEAEs by SOC, preferred term and maximum intensity, if a subject experiences more than 1 episode of a particular coded adverse event, the subject will be counted only once by the maximum intensity of the episode (preferred term). Similarly, if a subject has more than 1 adverse event within an SOC, the subject will be counted only once by the

maximum intensity in that SOC. Adverse events with missing severity will be classified as having the highest severity.

TEAEs classified in the eCRF as possibly or probably related to the study medication will also be summarized by preferred term and SOC. If a subject experiences more than 1 episode of a particular coded adverse event, the subject will be counted only once by the most related report for the preferred term. Similarly, if a subject has more than 1 adverse event within an SOC, the subject will be counted only once by the most related report in that SOC. Adverse events with missing relationship will be classified as having the highest relationship to study drug.

The following summaries will be presented:

- Overview of TEAEs during the study - number and percentage of subjects, number of events.
- TEAEs by SOC and preferred term - number and percentage of subjects.
- TEAEs by SOC and preferred term by sex (male and female) - number and percentage of subjects.
- TEAEs by SOC and preferred term by age - number and percentage of subjects.
- TEAEs by SOC - number and percentage of subjects.
- TEAEs by preferred term - number and percentage of subjects.
- Intensity of TEAEs by SOC and preferred term - number and percentage of subjects.
- Relationship of TEAEs by SOC and preferred term - number and percentage of subjects.
- Drug-related TEAEs by SOC and preferred term - number and percentage of subjects.
- TEAEs leading to study discontinuation by SOC and preferred term - number and percentage of subjects.
- Treatment-emergent Serious AEs by SOC and preferred term - number and percentage of subjects.
- Most Frequent(>=5%) Treatment-emergent Non-serious AEs by SOC and preferred term - number and percentage of subjects.
- Pretreatment Events by SOC and preferred term - number and percentage of subjects.

SOCs will be sorted in alphabetical order. Within an SOC, adverse events will be sorted in descending order of total number of subjects with the preferred term.

All adverse events will be presented in listing. Special listings for TEAEs leading to study discontinuation, SAEs and deaths will also be presented.

7.9.2 Vital Signs and Weight

Vital signs and weight at scheduled visits and their changes from Baseline will be summarized using descriptive statistics by visit and end of study. The number and percentage of subjects with

at least one post-Baseline MAV from vital signs during the treatment period will be presented for each variable over all visits. A listing of MAVs for vital signs will also be presented.

The criteria for identification of MAV vital signs are given in Appendix A.

7.9.3 Other Observations Related to Safety

The C-SSRS Baseline/Screening will be administered at Screening and the Since-Last-Visit C-SSRS will be administered at all other time points. The C-SSRS was developed by researchers at Columbia University as a tool to help systematically assess suicidal ideation and behavior in subjects during participation in a clinical trial of centrally acting drugs. The C-SSRS is composed of 3 questions addressing suicidal behavior and 5 questions addressing suicidal ideation, with sub-questions assessing severity. The tool is administered via an interview with the subject. If possible, the same interviewer should be used throughout the study for the same subject. It takes approximately 10 minutes to complete the C-SSRS.

7.10 Interim Analysis

The responder rate based on GAS score >=50 will be calculated. The conditional probability of a responder rate of less than 50% at 120 subjects will be assessed; that is, using the data /trend of 60 patients at the interim look, the probability of (responder rate < 50%) at the end of study (i.e., 120 subjects) will be calculated.

In addition, the change in PHQ9 and PDQ will be evaluated for internal decision making.

The report to be prepared for the interim analysis should include the tables for the following: demographics, disposition, GAS Response percents, GAS change from baseline, PHQ-9, and PDQ-D.

The end of the study will be the date of the last safety follow-up visit of the last subject based either on the first 60 completed subjects or the final 120 completed subjects.

7.11 Changes in the Statistical Analysis Plan

To be identified.

8.0 REFERENCES

Turner-Stokes, Lynne; *Goal Attainment Scaling (GAS) in Rehabilitation: A Practical Guide*. Clinical Rehabilitation 2009, 23: 362-370.

Appendix A. Criteria for Identification of Markedly Abnormal Values for Vital Signs

Parameter	Unit	MAV Definition (a)
Systolic Arterial BP	mmHg	< 85 mmHg or >180 mmHg
Diastolic Arterial BP	mmHg	<50 mmHg or >110 mmHg
Pulse	beats per minute	<50 bpm or >120 bpm
Weight	kg	Change of ≥7% body weight

BP=Blood pressure.
(a) MAV criteria are applied to postbaseline values and changes relative to Baseline values, as appropriate.