| | |
|---|---|
| Sponsor | Emalex Biosciences, Inc. |
| Protocol Title: | A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Parallel-Group, Phase 2 Exploratory Study to Evaluate the Efficacy and Safety of Ecopipam Tablets in Adults with Childhood Onset Fluency Disorder (Stuttering) |
| Protocol Number: | EBS-101-COFD-201 |
| Premier Research PCN: | EMAL210851 |
| Document Version: | Final 2.0 |
| Document Date: | 11-Mar-2022 |
| NCT Number: | NCT04492956 |

## Approvals

| Role | Signatures | Date (dd-Mmm-yyyy) |
|---|---|---|
| Biostatistician | Print Name: Kristin Wilborn<br>Sign Name: | |
| Premier Senior Statistical Reviewer | Print Name: Whitney Stemple<br>Sign Name: | |

| Role | Signatures | Date (dd-Mmm-yyyy) |
|---|---|---|
| Emalex Biosciences, Inc. Consultant | Print Name: Kai Fai Ho<br>Sign Name: | |
| Emalex Biosciences, Inc. Representative | Print Name: Kerensa Saljooqi<br>Sign Name: | |
| Emalex Biosciences, Inc. Representative | Print Name: Atul R. Mahableshwarkar<br>Sign Name: | |

## Document History

- Updated to add two shift tables (14.3.5.2.1 - Laboratory Data: Shift from Baseline in Chemistry Parameters by Study Visit and 14.3.5.7.1 – Laboratory Data: Shift from Baseline in HbA1c by Study Visit) and additional figures (SSS, CGI-I, PHQ-9, STAI-State)
- Whitney Stemple added as new senior reviewer

Table of Contents

## List of Tables



## List of Figures

# 1. Overview

This statistical analysis plan (SAP) describes the planned analysis and reporting for Emalex Biosciences, Inc. protocol number EBS-101-COFD-201 (A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Parallel-Group, Phase 2 Exploratory Study to Evaluate the Efficacy and Safety of Ecopipam Tablets in Adults with Childhood Onset Fluency Disorder (Stuttering), dated 16-Jul-2021 and protocol version 4.0. Reference materials for this statistical plan include the protocol and the accompanying sample data collection documents. Operational aspects related to collection and timing of planned clinical assessments are not repeated in this SAP unless relevant to the planned analysis.

The structure and content of this SAP provides sufficient detail to meet the requirements identified by the Food and Drug Administration (FDA), European Medicines Agency (EMA), and International Conference on Harmonization (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use: Guidance on Statistical Principles in Clinical Trials[1]. All work planned and reported for this SAP will follow internationally accepted guidelines, published by the American Statistical Association[2] and the Royal Statistical Society[3], for statistical practice.

The planned analyses identified in this SAP may be included in clinical study reports (CSRs), regulatory submissions, or future manuscripts. Also, post-hoc exploratory analyses not necessarily identified in this SAP may be performed to further examine study data. Any post-hoc or unplanned, exploratory analysis performed will be clearly identified as such in the final CSR.

The statistical plan described hereafter is an *a priori* plan. It will be submitted to file prior to any unblinded inferential or descriptive analysis of data pertaining to Emalex Biosciences, Inc.'s study EBS-101-COFD-201.

# 2. Study Objectives and Endpoints

## 2.1. Study Objectives

### 2.1.1. Primary Objective

The primary objective is to evaluate the efficacy of ecopipam tablets in adult subjects with childhood onset fluency disorder.

### 2.1.2. Secondary Objectives

The secondary objectives are to evaluate the safety of ecopipam tablets in adult subjects with childhood onset fluency disorder and to characterize the pharmacokinetics (PK) of ecopipam and its active metabolite (EBS-101-40853) after oral administration of ecopipam in this patient population.

### 2.1.3. Exploratory Objectives

To evaluate the potential PK/pharmacodynamic relationship between plasma levels of ecopipam and EBS-101-40853 and safety and/or treatment effect, if possible.

## 2.2. Study Endpoints

### 2.2.1. Efficacy Endpoints

#### 2.2.1.1. Primary Efficacy Endpoint

The primary efficacy endpoint of this study is the change in the Stuttering Severity Instrument 4th Edition (SSI-4) from Baseline to Week 12 for subjects receiving ecopipam compared with subjects receiving placebo as assessed by a central rater.

#### 2.2.1.2. Key Secondary Efficacy Endpoints

The key secondary efficacy endpoints are:

- Change from Baseline in the Subjective Screening of Stuttering (SSS) score for subjects receiving ecopipam compared with subjects receiving placebo from Baseline to Week 12.
- Improvement in the Clinical Global Impression-Severity (CGI-S) score for subjects receiving ecopipam compared with subjects receiving placebo from Baseline to Week 12.

#### 2.2.1.3. **Other S**econdary Efficacy Endpoints

The secondary efficacy endpoints of this study include the following:

- Change in the Overall Assessment of the Speaker's Experience of Stuttering (OASES) score for subjects receiving ecopipam compared with subjects receiving placebo from Baseline to Week 12.
- Clinical Global Impression – Improvement (CGI-I) score for subjects receiving ecopipam compared with subjects receiving placebo at Week 12.
- Change in the SSI-4 percentage of syllables stuttered for subjects receiving ecopipam compared with subjects receiving placebo from Baseline to Week 12.
- Change in the SSI-4 frequency sub-score for subjects receiving ecopipam compared with subjects receiving placebo from Baseline to Week 12.
- Change in the SSI-4 duration sub-score for subjects receiving ecopipam compared with subjects receiving placebo from Baseline to Week 12.
- Change in the OASES Section IV score for subjects receiving ecopipam compared with subjects receiving placebo from Baseline to Week 12.
- Changes in all primary and secondary endpoints for subjects receiving ecopipam compared with subjects receiving placebo from Baseline to Weeks 4 and 8.

### 2.2.2. Safety Endpoints

The safety endpoints of this study include the following:

- All adverse events (AEs) and serious adverse events (SAEs) (all visits)
- Hematology, blood chemistry, and urine values at Screening, Baseline, Week 12, and 7 and 14 day Follow-up visits

- HbA1c (Baseline and Week 12)
- Measurements of vital signs (height, weight, orthostatic blood pressure, and pulse) at Baseline, Weeks 4, 8 and 12, and 7 and 14 day Follow Up visits
- 12-lead electrocardiogram (ECG) at Baseline, Weeks 4, 8 and 12, and 7 and 14 day Follow Up visits
- Complete physical examination at Baseline and Week 12 and brief physical exam as necessary per PI judgment based on reported symptoms at all other clinic visits
- Columbia-Suicide Severity Rating Scale (C-SSRS) at all visits except 30 day Follow Up visit)
- Patient Health Questionnaire-9 (PHQ-9) at Screening, Baseline and Weeks 4, 8, and 12
- State-Trait Anxiety Inventory for Adults™ (STAI-AD) at Baseline and Weeks 4, 8, and 12)
- Extra-pyramidal Symptom Rating Scale (ESRS) at Baseline and Weeks 4, 8, and 12
- Signs or symptoms or withdrawal, abuse, and dependence throughout the study

#### 2.2.3. Exploratory Pharmacokinetic/Pharmacodynamic Variables

The PK endpoints of the study include the concentrations of ecopipam and its major (active) metabolite, EBS-101-40853, at Weeks 4 and 12.

## 3. Overall Study Design and Plan

### 3.1. Overall Design

This is a multicenter, randomized, double-blind, placebo-controlled, parallel-group, Phase 2 exploratory study in adult subjects with childhood onset fluency disorder.

The Central Enrollment Committee (CEC) has been developed as a quality measure to focus on ensuring that all of the inclusion are met and none of the exclusion criteria are met, and confirmation of diagnosis based on the comprehensive information gathered and verified by the Investigator. Subjects who receive CEC affirmation during the 28-day screening period may proceed to the Baseline visit.

At the Baseline visit, eligible subjects will be randomized 1:1 to receive either a target steady-state dose of ecopipam ~2 mg/kg/day or matching placebo for a 12-week Treatment Period consisting of a 4-week Titration Phase followed by an 8-week Maintenance Phase.

Subjects will return to the clinic 4, 8, and 12 weeks after Randomization and Follow-up visits 7 and 14 days after completing the Treatment Period or Early Discontinuation. Efficacy assessments will be conducted at Weeks 4, 8, and 12, and safety assessments will be conducted at all visits. Subjects will have AEs and other safety parameters assessed by phone or video conference at Weeks 2, 6, and 10, and 30 days after the last study drug administration. Signs or symptoms of withdrawal, abuse, and dependence will be monitored throughout the study.

The SSI-4 will be administered by a dedicated site rater who is trained and certified by the Sponsor-approved central rater. All SSI-4 interviews will be video recorded.

At the end of the Treatment Period or Early Discontinuation, subjects will taper study drug by 25 mg/day until off study drug, for up to 1 week.

Where possible, safety and efficacy assessments will be made available for remote administration in the case that restrictions, as a consequence of the novel coronavirus disease-2019 (COVID-19) or other natural disaster, prevent in-clinic study visits.

### 3.2. Sample Size and Power

A sample size of 68 subjects (34 subjects per treatment arm) will provide ≥80% power to detect a 20% difference on the SSI-4 between ecopipam and placebo groups on change from Baseline to Week 12, with standard deviation = 24.7 and alpha = 0.05, assuming 10% drop out.

### 3.3. Study Population

This study will enroll male and female adult subjects (≥ 18 years of age), weighing ≥ 45 kg, with childhood onset (prior to 8 years old) fluency disorder.

### 3.4. Treatments Administered

Ecopipam or placebo 12.5, 50, and 75 mg tablets will be supplied for a target dose of ~2 mg/kg/day during the 4-week Titration Phase and the 8-week Maintenance Phase for each of the weight bands. All doses will be administered once daily in the evening at bedtime, except for in-clinic dosing at Visit 4.

At the end of the Treatment Period or Early Discontinuation, subjects will taper study drug by 25 mg/day until off study drug, for up to 1 week.

Dosing will be stratified by the following weight bands to better achieve the target dose: 45 to ≤68 kg; 69 to ≤83 kg; and >83 kg.

**Table 1: Dosing Weight Bands**

| | Titration Phase | | | | Maintenance Phase |
|---|---|---|---|---|---|
| **Weight (kg)** | **Dose (mg) Week #1** | **Dose (mg) Week #2** | **Dose (mg) Week #3** | **Dose (mg) Week #4** | **Dose (mg) Week #5 - 12** |
| 45 to ≤68 | 25 (2 × 12.5) | 50 | 75 | 100 (2 x 50) | 100 (2 x 50) |
| 69 to ≤83 | 25 (2 × 12.5) | 50 | 100 (2 x 50) | 150 (2 × 75) | 150 (2 × 75) |
| >83 | 25 (2 × 12.5) | 50 | 100 (2 x 50) | 200 (4 x 50) | 200 (4 x 50) |

### 3.5. Method of Assigning Subjects to Treatment Groups

Subjects will be randomized 1:1, stratified by weight band, to receive either a target steady-state ecopipam HCl dose of ~2 mg/kg/day or matching placebo.

### 3.6. Blinding and Unblinding

The subject and all personnel involved with the conduct and interpretation of the study, including the Investigators, study site personnel, the Sponsor and contract research organization (CRO) clinical staff, including the Medical Monitor, will be blinded to study medication

### 3.7. Schedule of Events

A detailed schedule of events for the study is provided in Table 2.

## Table 2: Schedule of Events

| Assessment | Screening | Baseline | Treatment Period | | | | | Taper Phase# Completion/ Early Termination | Follow-Up | | Unscheduled∞ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Titration Phase | | Maintenance Phase | | | | Day 7 & 14 | Day 30 | |
| **Study Week** | **Up to -4** | **0** | **2*** | **4** | **6*** | **8** | **10*** | **12** | **13 & 14** | **16*** | |
| **Study Days** | **Up to -28** | **1** | **14±3** | **28±3** | **42±3** | **56±3** | **70±3** | **84±3** | **91±3, 98±3** | **114±3** | |
| **Study Visit Number** | **V1** | **V2** | **V3** | **V4** | **V5** | **V6** | **V7** | **V8** | **V9 & V10** | **V11** | |
| Informed Consent[a] | X | | | | | | | | | | |
| Inclusion/ Exclusion | X | X | | | | | | | | | |
| Medical/Psychiatric/ Medication History | X | X | | | | | | | | | |
| CEC eligibility determination | X | | | | | | | | | | |
| Randomization | | X | | | | | | | | | |
| Physical Exam/Vital Signs[b] | X | X | | X | | X | | X | X | | X |
| Local ECG | X | X | | X | | X | | X | X | | X |
| Laboratory tests (Hematology, Chemistry and urinalysis)[c] | X | X | | X | | X | | X | X | | X |
| Urine Drug Screen | X | X | | | | | | X | | | X |

| Assessment | Screening | Baseline | Treatment Period | | | | | Taper Phase# Completion/ Early Termination | Follow-Up | | Unscheduled∞ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Titration Phase | | Maintenance Phase | | | | Day 7 & 14 | Day 30 | |
| **Study Week** | **Up to -4** | **0** | **2*** | **4** | **6*** | **8** | **10*** | **12** | **13 & 14** | **16** | |
| **Study Days** | **Up to -28** | **1** | **14±3** | **28±3** | **42±3** | **56±3** | **70±3** | **84±3** | **91±3, 98±3** | **114±3** | |
| **Study Visit Number** | **V1** | **V2** | **V3** | **V4** | **V5** | **V6** | **V7** | **V8** | **V9 & V10** | **V11** | |
| Urine Pregnancy Test | X | X | | | | | | X | | | X |
| DSM-5 Criteria for Stuttering checklist | X | | | | | | | | | | |
| SSI-4[d] | X | X | | X | | X | | X | | | X |
| CGI-S | | X | | X | | X | | X | | | X |
| CGI-I | | | | X | | X | | X | | | X |
| SSS | | X | | X | | X | | X | | | |
| OASES | | X | | X | | X | | X | | | |
| C-SSRS | X | X | X | X | X | X | X | X | X | | X |
| ESRS | | X | | X | | X | | X | | | |
| PHQ-9 | X | X | | X | | X | | X | | | X |
| STAI-AD | | X[e] | | X | | X | | X | | | X |
| PK Blood Draws | | | | X[f] | | | | X[g] | | | |

| Assessment | Screening | Baseline | Treatment Period | | | | | Taper Phase# Completion/ Early Termination | Follow-Up | | Unscheduled∞ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Titration Phase | | Maintenance Phase | | | | Day 7 & 14 | Day 30 | |
| **Study Week** | **Up to -4** | **0** | **2*** | **4** | **6*** | **8** | **10*** | **12** | **13 & 14** | **16*** | |
| **Study Days** | **Up to -28** | **1** | **14±3** | **28±3** | **42±3** | **56±3** | **70±3** | **84±3** | **91±3, 98±3** | **114±3** | |
| **Study Visit Number** | **V1** | **V2** | **V3** | **V4** | **V5** | **V6** | **V7** | **V8** | **V9 & V10** | **V11** | |
| Assess Adverse Events | X | X | X | X | X | X | X | X | X | X | X |
| Assess Concommitant Medications | X | X | | X | X | X | | X | X | | X |
| Dispense Study Drug[a] | | X | | X[f] | | X | | X | | | X |
| Collect Unused Study Drug/Assess drug compliance | | | | | | | | | | | |

Abbreviations: AE = adverse event; BP = blood pressure; CEC = Central Enrollment Committee; CGI-S/I = Clinical Global Improvement-Severity/Improvement; C-SSRS = Columbia-Suicide Severity Rating Scale; DSM-5 = Diagnostic and Statistical Manual for Mental Disorders, 5th edition; ECG = electrocardiogram; ESRS = Extra-pyramidal Symptom Rating Scale; HR = heart rate; OASES = Overall Assessment of the Speaker's Experience of Stuttering; PHQ-9 = Patient Health Questionnaire-9; PK = pharmacokinetic(s); SSI-4 = Stuttering Severity Instrument; SSS = Subjective Screening of Stuttering; STAI-AD = State-Trait Anxiety Inventory for Adults;

[a] Informed consent must be performed prior to screening assessments.

[b] A complete physical examination will be conducted at Baseline and Week 12 and a brief physical exam as necessary per PI judgment based on reported symptoms at all other clinic visits. Vital signs will include, pulse, HR, BP, orthostatic BP and pulse (done 5 minutes after being supine and then 3 minutes after standing), weight and height (height measured at Screening only).

[c] Subjects should be in a fasting state (8 hours) for laboratory tests after Visit 1 Screening. HbA1c will be measured at the Baseline and Completion/Early Termination visits.

[d] The SSI-4 will be administered by a Sponsor-approved, qualified rater. All SSI-4 interviews are conducted via video recording with the subject and interviewer in separate locations and recordings will be scored by the central rater. The central rater scores will be used for all analyses. Every attempt should be made, per subject, to perform the SSI-4 interview by the same SSI-4 interviewer at the same time and same location at each visit.

[e] The STAI form Y-1 (State) will be completed at all marked visits. The STAI form Y-2 (Trait) will be completed at Baseline visit only and administered **after** the State portion (STAI Form Y-1) as per STAI-AD Manual (page 13).

[f] A morning appointment will be made and subjects will be instructed not to take their prescribed dose of study drug within the previous 24 hours. Subjects

will be instructed to record the date and time of the last administration of study drug prior to the study visit. Three PK samples will be collected at 1) pre-dose (20 to 36 hours since the last dose); and 2) post-dose between 0.5 and 1.5 hours; and 3) post-dose between 2 and 4 hours after study drug administration. Any samples should be collected at least 30 min apart. A missed PK collection due to consequences of COVID-19 or other natural disaster will not be considered a protocol deviation.

g Subjects will be instructed to record the date and time of the last administration of study drug prior to the study visit. A PK blood sample will be collected at the end of the visit with the time of sample collection recorded. A missed PK collection due to consequences of COVID-19 or other natural disaster will not be considered a protocol deviation.

h Site personnel who are delegated the responsibility to dispense tudy drug may coordinate the shipment of study drug to be delivered to a subject's residence by in the case that restrictions, as a consequence of COVID-19 or other natural disaster, prevent in-clinic study visits.

# Study drug will be dispensed for the Taper Phase reducing by 25 mg/day until off study drug.

* Weeks 2, 6, 10, and 16 are remote assessment visits. If there are any abnormal findings, the subject may be brought to the site for a full assessment. The remote assessment visit at Week 16 (30 days after the last study drug administration) is to assess any adverse events that may have occurred.

∞ PI will discuss which assessments are best suited for the Unscheduled visit with Study Medical Monitor.

µ Where possible safety and efficacy assessments will be made available for remote administration in the case that restrictions, as a consequence of COVID-19 or other natural disaster, prevent in-clinic study visits.

# 4. Statistical Analysis and Reporting

## 4.1. Introduction

Data processing, tabulation of descriptive statistics, calculation of inferential statistics, and graphical representations will be performed primarily using SAS (release 9.4 or higher). If the use of other software is warranted, the final statistical methodology report will detail what software was used for what purposes.

Continuous (quantitative) variable summaries will include the number of subjects (n) with non-missing values, mean, standard deviation (SD) median, minimum, and maximum.

Categorical (qualitative) variable summaries will include the frequency and percentage of subjects who are in the particular category or each possible value. In general, the denominator for the percentage calculation will be based upon the total number of subjects in the study population for the treatment groups, unless otherwise specified. The denominator for by-visit displays will be the number of subjects in the relevant study population with non-missing data at each visit.

The minimum and maximum will be reported with the same degree of precision (i.e., the same number of decimal places) as the observed data. Measures of location (mean and median) will be reported to 1 degree of precision more than the observed data and measures of spread (SD) will be reported to 2 degrees of precision more than the observed data.

Percentages will be presented to 1 decimal place, unless otherwise specified.

Unless otherwise indicated, all statistical tests will be conducted of $\leq 0.05$ will be considered statistically significant using 2-tailed tests. *P* values and corresponding 95% confidence intervals (CIs) will be presented for statistical tests.

Statistical testing will be performed at the 0.05 level using two-tailed tests. A *P* value of $\leq 0.10$ but $> 0.05$ will be considered evidence of a trend.

## 4.2. Interim Analysis and Data Monitoring

No interim analyses are planned.

# 5. Analysis Populations

The following analysis populations are planned for this study:

- **Safety Population (SAF):** The Safety Population includes all subjects who received at least 1 dose of study drug. The Safety Population will be used for the analysis of the safety endpoints. Subjects will be analyzed by actual treatment received.
- **Modified Intent-To-Treat Population (mITT):** The mITT population includes all randomized subjects who received at least 1 dose of study drug and have at least 1 post-baseline efficacy evaluation completed. The mITT population will be used for the analysis of primary and secondary efficacy endpoints. Subjects will be analyzed by randomized treatment.

- **Pharmacokinetic Population (PK):** The PK Population includes all subjects from the Safety population who have a valid concentration measurement.
- **Per Protocol (PP)**: The PP Population includes all subjects from the mITT population who have no important protocol deviations. Subjects will be analyzed by randomized treatment.

  Before data are released for statistical analysis, a blinded review of all data will be performed by the Sponsor's clinical team to identify protocol deviations that may potentially affect the results. At that time, it will be determined if subjects should be excluded from the PP Population. The list of subjects to be excluded from the PP Population, along with the reason for exclusion, will be finalized prior to database unblinding.

# 6. General Issues for Statistical Analysis

## 6.1. Statistical Definitions and Algorithms

### 6.1.1. Baseline

The last non-missing observation recorded before receiving the first dose of study drug will be used as the baseline observation for all calculations of change from baseline.

### 6.1.2. Adjustments for Covariates

If there is a statistical difference among treatment groups with respect to baseline characteristics, that variable may be added to the statistical models as a covariate to determine the effect on treatment. Baseline efficacy scale scores and study site are planned for inclusion in statistical models, unless otherwise indicated. See Section 8 for more details.

### 6.1.3. Multiple Comparisons

The type I error rate for the study will be preserved by using sequential (hierarchical) testing for the primary, key secondary, and secondary endpoints of the mITT population. The key secondary endpoint will only be tested if ecopipam treatment is significantly different from placebo for the primary endpoint.

The hierarchical order for testing the null hypotheses will be done in the following order:

Primary Endpoint:

1. Mean change from baseline in SSI-4 total score at Week 12.

Key Secondary Endpoints:

2. Mean change from baseline in SSS at Week 12.
3. Mean change from baseline in CGI-S score at Week 12.

Other Secondary Endpoints:

4. Mean change from baseline in OASES overall impact score at Week 12.
5. Summary of CGI-I at Week 12.

6. Mean change from baseline in the SSI-4 percentage of syllables stuttered at Week 12.
7. Mean change from baseline in the SSI-4 frequency subscore at Week 12.
8. Mean change from baseline in the SSI-4 duration subscore at Week 12.
9. Mean change from baseline in the OASES section IV (Quality of Life) subscore at Week 12.
10. Mean change from baseline in SSI-4 total score at Week 8.
11. Mean change from baseline in SSS at Week 8.
12. Mean change from baseline in CGI-S score at Week 8.
13. Mean change from baseline in OASES overall impact score at Week 8.
14. Summary of CGI-I at Week 8.
15. Mean change from baseline in the SSI-4 percentage of syllables stuttered at Week 8.
16. Mean change from baseline in the SSI-4 frequency subscore at Week 8.
17. Mean change from baseline in the SSI-4 duration subscore at Week 8.
18. Mean change from baseline in the OASES section IV (Quality of Life) sub-score at Week 8.
19. Mean change from baseline in SSI-4 total score at Week 4.
20. Mean change from baseline in SSS at Week 4.
21. Mean change from baseline in CGI-S score at Week 4.
22. Mean change from baseline in OASES overall impact score at Week 4.
23. Summary of CGI-I at Week 4.
24. Mean change from baseline in the SSI-4 percentage of syllables stuttered at Week 4.
25. Mean change from baseline in the SSI-4 frequency subscore at Week 4.
26. Mean change from baseline in the SSI-4 duration subscore at Week 4.
27. Mean change from baseline in the OASES section IV (Quality of Life) subscore at Week 4.

Because the testing is sequential, the type I error rate of $\alpha = 0.05$ is maintained. Failure at any stage in the sequence implies automatic failure at all subsequent stages. Those tests will be considered as exploratory without any formal conclusions drawn from them. All *P* values for each comparison without adjustments will be provided in the summary tables for informational purposes.

No adjustments will be made for multiple comparisons for other endpoints.

### 6.1.4. Handling of Dropouts or Missing Data

All possible efforts will be made to ensure that subjects stay in the study and all data is collected as scheduled however, the occurrence of missing data cannot be completely eliminated. The primary analysis population for all efficacy analyses will be the mITT Population. Several methods will be used to handle missing data for the primary analysis only (SSI-4 total score), including:

- MMRM
- MI using placebo-based (jump to control) imputation

The MMRM method with a Kenward-Roger approximation to estimate degrees of freedom for the denominator will be used for the primary and secondary analyses. By default, MMRM

implements a standard MAR approach using the restricted maximum likelihood (REML) method. A sensitivity analysis using the MI method with placebo-based imputation will be performed to assess the robustness of the MAR assumptions for the primary endpoint only for the mITT population.

The above procedures are described in the below sub-sections.

Any subject who withdraws from the study or who misses a scheduled visit or assessment through Week 12 will have their primary efficacy data (SSI-4 total score) analyzed as observed (with missing data not imputed) and will also have their efficacy data analyzed as imputed using multiple imputation, which will be presented as a sensitivity analysis. Secondary efficacy data will be analyzed as observed (with missing data not imputed). Safety data will not be imputed and will be analyzed as observed.

PK data at Weeks 4 and 12 may not be collected due to COVID-19. Any PK data not collected at these time points will be notated in listings and not analyzed in tables.

#### 6.1.4.1. MMRM Analysis Method

The SSI-4 total score will be analyzed using a MMRM model as the primary analysis method, with change from baseline as the dependent variable and fixed, categorical effects of treatment group, visit, study site, weight-band, and treatment group-by-visit interactions as well as the continuous, fixed covariate of baseline SSI-4 total score.

The MMRM method has been demonstrated extensively as an appropriate choice for the primary analysis in longitudinal confirmatory clinical trials with continuous endpoints (Mallinckrodt et. Al., 2008[4]). This analysis method, which is from a broader class of direct-likelihood analyses methods, makes use of fully and partially observed data sequences from individual subjects by estimating the covariance between data from different time points (Molenberghs and Kenward, 2007[5]). Further, it is often useful to implement MMRM using an unstructured approach to modeling both the treatment-by-time means and the variances and covariances, leading to what is essentially a multivariate normal model wherein treatment group means at the primary time point are adjusted to reflect both the actual observed data and the projected outcomes from the subjects with missing data (Cnaan et al., 1997[6], Molenberghs et al., 2004[7], Molenberghs and Kenward, 2007[5]).

As a direct likelihood method, the MMRM method is a preferred approach for handling missing data in such designs and will be used as the primary analysis method for all efficacy endpoints where assessments are made over time. MMRM is a full multivariate model in nature, which avoids potential bias as a predetermined model and operates in a more general missing at random (MAR) framework (Mallinckrodt et al., 2001[8]). Data are considered MAR if, conditional upon the independent variables in the analytic model, the missingness depends on the observed outcomes of the variable being analyzed but does not depend on the unobserved outcomes of the variable being analyzed. This assumption implies that the behavior of the post dropout observations can be predicted from the observed variables, and therefore that treatment effect can be estimated without bias using the observed data (European Medicines Agency, 2010[9]). For studies of missing data in a controlled clinical trial setting, MAR is usually considered as a plausible underlying missing mechanism (Molenberghs and Kenward, 20075; Siddiqui et al., 200910, Mallinckrodt et al., 2008[4], Mallinckrodt et al., 2013[11]). The assumption of MAR is often

reasonable because, particularly in longitudinal studies wherein the evolution of treatment effects is assessed by design over time, the observed data and the models used to analyze them can explain much of the missingness (Little and Rubin, 1987[12], Verbeke and Molenberghs, 2000[13]). This point may be especially relevant in well-controlled studies, in which extensive efforts are made to observe all outcomes and factors that influence them while subjects are following protocol-defined procedures. Thus, longitudinal clinical trials by their very design aim to reduce the amount of missing not at random (MNAR) data (missingness explained by unobserved responses), thereby increasing the plausibility of MAR (Mallinckrodt et al., 2008[4]).

#### 6.1.4.2. Multiple Imputation Methods

Although the assumption of MAR, as used for the primary efficacy analysis method, is often reasonable in clinical trials, the possibility of MNAR data cannot be ruled out. Therefore, analysis valid under MNAR will also be performed. Both a MNAR- and MAR-based analyses using MI methods will be the basis upon which sensitivity of the analysis to missing data are assessed.

Any subject who withdraws or is discontinued from the study or who misses a scheduled visit or assessments up through Week 12 will have their primary efficacy missing data analyzed as imputed using MI techniques. This analysis will be presented as a sensitivity analysis.

Multiple imputation is a simulation-based approach where missing values are replaced using an appropriate stochastic model given the observed data and covariates, creating multiple completed data sets. These completed datasets are then analyzed using standard analysis methods (MMRM for this study), and the different parameter estimates across the datasets are then combined to produce unique point estimates, standard errors, and confidence intervals taking into account the uncertainty of the imputation process.

In most randomized clinical trials that collect data over time, the great majority of missing data follow a monotone pattern. That is, once a subject has a missing data for some visit, data will be missing for all subsequent visits. Typically, there is also a small amount of non-monotone missing data (i.e., some subjects have missing values for intermediate visits, but have non-missing data at subsequent visits).

The following MI analysis models, based on the MNAR approach, will be used to examine robustness of the primary analysis results.

MI with Placebo-Based Imputation

A placebo-based (jump to control) multiple imputation for missing SSI-4 total scores will be carried out for subjects who withdraw from the study or have missing data at a scheduled visit through Week 12, as indicated previously. The imputation has three broad components: i) the multiple imputation process for the placebo data; ii) the multiple imputation process for the ecopipam data; and iii) the analysis model that will be used to draw inference regarding the primary causal estimands along with the method for combining the results across the multiply-imputed datasets. The primary causal estimands are defined in Section 8.1.

*Step 1*: If the data has a non-monotone pattern of missingness, then a monotone data augmentation method using Markov-Chain Monte-Carlo (MCMC) will be used to impute data that is missing and required to make the missing data pattern monotone. Fifty datasets with a monotonic missing pattern will be generated. This method will use a non-informative Jeffreys prior to derive the posterior mode from the expectation-maximization (EM) algorithm as the starting values for the MCMC method. Intermittent missing values will be imputed using the MCMC method assuming a multivariate normal distribution over all variables included in the imputation model (i.e., treatment group, baseline, and each post-baseline visit). The MCMC statement of the MI procedure in SAS (PROC MI) will specify the CHAIN=MULTIPLE option, so that the procedure uses multiple chains and completes the default 200 burn-in iterations before each imputation, and the IMPUTE=MONOTONE option to create the 50 partially imputed datasets with a monotone missing pattern. The seed of the pseudorandom number generator used to randomly generate imputations for the missing values in Step 1 is 3590145.

Assumptions underlying the partial imputation step are such that subjects with missing data follow the same model as other subjects in their respective treatment arm that have complete data.

If the raw data has a monotone pattern of missingness, then the same procedures described above can be followed to create 50 identical datasets that will be used as an input dataset for the next step.

*Step 2*: The second stage will impute missing post-informative subject withdrawal monotone data in a sequential manner, using a method proposed by Ratitch and O'Kelly[14]. The post-informative subject withdrawal values in each of the imputed datasets for the ecopipam arm will be set to missing, thus leaving a dataset with multiply-imputed values and monotone missing data patterns due to early termination as a result of AEs, lack of efficacy, i.e., this dataset will contain the complete data for all placebo subjects (observed and imputed under MAR) and the ecopipam subjects (intermittent missing data imputed under MAR and monotone missing data for non-informative withdrawal imputed under MAR), with monotone missing data due to informative withdrawal left as missing (to be imputed under MNAR). The second stage of the imputation procedure will handle the MNAR imputation for the monotone missing post-informative subject withdrawal data in the ecopipam subjects. A new dataset will then be created with all placebo subjects including their observed and imputed data, as well as records from ecopipam that are classified as MNAR. Then ecopipam subjects will have all MNAR missing data imputed based on placebo data. Once the MNAR data is imputed, this data (containing all placebo data and imputed MNAR data for ecopipam) will then be combined with non-missing and imputed MAR data from ecopipam subjects to create the datasets to be used for analysis. The second stage imputation will be done sequentially by study visit for each first stage imputation. The seed number of 9482038 will be used for the imputation procedure described in Step 2.

For both Steps 1 and 2, minimum and maximum values for the SSI-4 total score (15 and 135) will be specified in the MI procedure to avoid imputed values outside the possible range of values. When an intended imputed value is less than the minimum or greater than the maximum value specified, the MI procedure in SAS will redraw another value for imputation.

*Step 3*: MMRM analyses will be performed separately for each of the 50 complete analysis datasets, and the results will be combined into one multiple imputation inference (estimated treatment effect, standard error, p-value and associated 95% confidence interval) using the SAS MIANALYZE procedure. The treatment difference will be tested at the 2-sided 0.05 level and corresponding 95% CIs will be calculated.

Tipping Point Analysis

Results from the placebo-based MI analysis will be compared to the MMRM analysis. If these results differ (e.g., MI analysis is significant, MMRM is not), then an additional tipping point sensitivity analysis will be performed in order to determine the inflection point at which the inference under the MNAR assumption changes substantially. This will be used to check the robustness of the imputation.

The tipping point analysis will apply a shift parameter, which will adjust the imputed values for observations in the active treatment group. The range of shift parameter will be determined based on the resulting standard deviation of the change from baseline SSI-4 score at Week 12 for the group that received placebo. Once the likely point of the shift is determined, the analysis will be rerun using an expanded range around the suspected tipping point, with greater precision (i.e., by 1). Thus, the value at which the results of the analysis are shifted from significant (i.e., $\alpha \leq 0.05$) to non-significant (i.e., $\alpha > 0.05$) will be determined.

The seed number of 2740211 will be used for the imputation procedure described in Step 2.

Similar to the standard MAR imputation method, options for minimum and maximum values will be used to make the imputed values more consistent with the observed variable values. When an intended imputed value is less than the minimum or greater than the maximum value specified in the MI procedure, SAS will redraw another value for imputation.

Steps 1 and 3 of the analysis will be the same as for the multiple imputation analysis as described in in placebo-based imputation above. However, Step 2 of the analysis is where the shift parameters will be applied. Pseudo-code for Step 2 is as follows:

```
proc mi data=example_1 seed=2740211 nimpute=XX out=example_2;
  class <treatment>;
  monotone regpmm(<visit02 score> ..... <visit12 score>);
  var <treatment> <baseline score> ..... <visit12 score>;
  mnar adjust(<visit02 score> / shift=YY adjustobs=(treatment = 'ecopipam'));
  mnar adjust(<visit04 score> / shift=YY adjustobs=(treatment = 'ecopipam'));
  mnar adjust(<visit08 score> / shift=YY adjustobs=(treatment = 'ecopipam'));
  mnar adjust(<visit12 score> / shift=YY adjustobs=(treatment = 'ecopipam'));
run;

XX will be the determined based on the proportion of missing data across visits.
YY will encompass the range of shift parameters as specified above.
```

### 6.1.5. Analysis Visit Windows

Analyses of all efficacy variables for this study will be done using the visit windows outlined below. Safety data will not be windowed and will be analyzed as per the nominal visit present in the database.

Efficacy data collected at in-clinic study visits, including the CGI-S, CGI-I, SSI-4, SSS, and OASES, will be assigned to a study visit for analyses based on the study day relative to the visit according to Table 3:

**Table 3: Efficacy Variable Visit Windows**

| Study Visit | Target Day | Study Day Intervals |
|---|---|---|
| Week 4 | 28 | Day 15 to 42 |
| Week 8 | 56 | Day 43 to 70 |
| Week 12 | 84 | Day 71 to EOT |
| EOT = end of treatment | | |

For those subjects who discontinue early from the study, Table 3 will also be used to assign the appropriate analysis visit.

The study day will be calculated for each scheduled or EOT post-baseline visit and compared to the assessment window presented in Table 3 to define the visit window used for analyses. The analysis visit windows only apply to those visits that are applicable to the specific assessment.

If more than 1 assessment occurs within a single visit window, then the analysis will use the assessment closest to the target day. If 2 assessments within the same visit window are equidistant from the target day, then the analysis will use the later assessment.

Efficacy assessments performed >14 days after the last dose of study medication will not be included in any efficacy analyses.

### 6.1.6. Pooling of Sites

Sites will be pooled for statistical analysis as follows. For analysis, each site is expected to have approximately 10 randomized subjects. The smallest sites will be grouped sequentially in order of smallest to largest, until each pooled site has a minimum of 10 subjects. If a site had 2 site numbers due to additional randomization numbers required (if needed), those site numbers will be considered as 1 site before pooling.

### 6.1.7. Derived Variables

- **Change from baseline** = value at current time point – value at baseline.
- **Date of last dose** =

- If the taper kit was dispensed and the number of tablets taken is known, then the **date of last dose** = date of first dose from the taper kit + # days at least 1 tablet taken – 1, where # days at least 1 tablet taken is derived as follows:
  - if 3 taper bottles were dispensed, then 1-3 tablets taken = 1 day; 4 tablets taken = 2 days; 5-6 tablets taken = 3 days
  - if 5 taper bottles were dispensed, then 1-2 tablets taken = 1 day; 3-4 tablets taken = 2 days; 5 tablets taken = 3 days; 6 tablets taken = 4 days; 7-8 tablets taken = 5 days
  - if 7 taper bottles were dispensed, then 1-3 tablets taken = 1 day; 4-6 tablets taken = 2 days; 7-10 tablets taken = 3 days; 11-12 tablets taken = 4 days; 13 tablets taken = 5 days; 14 tablets taken = 6 days; 15-16 tablets taken = 7 days
- If the taper kit was dispensed and the number of tablets taken is not known (e.g., subject lost to follow-up), then the **date of last dose** = date of first dose from the taper kit + number of taper bottles dispensed – 1.
- Else if the taper kit was *not* dispensed *and* the subject discontinued early but was not lost to follow-up, then the **date of last dose** = (i) date last dispensed kit was returned – 1 day; OR if last dispensed kit was not returned, (ii) date of first dose of the last kit that was dispensed + 28 days
- Else if the taper kit was *not* dispensed *and* the subject was lost to follow-up, then the **date of last dose** = the earliest of (i) lost to follow-up date; OR (ii) date of first dose of the last kit that was dispensed + 28 days

- **Treatment emergent AE (TEAE)** = any AE that is newly occurring or worsening with an onset date/time after the first dose of study drug.

- **OASES I (General Information)** = sum of questions 1-20
- **OASES II (Reactions to Stuttering)** = sum of questions 21-50
- **OASES III (Communication in Daily Situations)** = sum of questions 51-75
- **OASES IV (Quality of Life)** = sum of questions 76-100
- **OASES Overall Score** = sum of each of the four sub-sections
- **OASES Impact Score** = Calculate an impact score for each of the sections and overall by dividing the total points by the number of items completed.
  - Mild = 1.00-1.49
  - Mild to Moderate = 1.50-2.24
  - Moderate = 2.25-2.99
  - Moderate to Severe = 3.00-3.74
  - Severe = 3.75-5.00
- **C-SSRS Suicidal Ideation** = A "yes" answer at any time during treatment to any 1 of the 5 suicidal ideation questions (Categories 1-5: Wish to be Dead, Nonspecific Active

Suicidal Thoughts, Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act, Active Suicidal Ideation with Some Intent to Act, without Specific Plan, and Active Suicidal Ideation with Specific Plan and Intent).

- **C-SSRS Suicidal Behavior** = A "yes" answer at any time during treatment to any 1 of the 5 suicidal behavior questions (Categories 6-10: Preparatory Acts or Behavior, Aborted Attempt, Interrupted Attempt, Actual Attempt (nonfatal), Completed Suicide).

- **C-SSRS Suicidal Ideation or Behavior** = A "yes" answer at any time during treatment to any 1 of the 10 suicidal ideation and behavior questions (Categories 1-10) on the C-SSRS.

- **PHQ-9** = sum of the 9 questions (individual questions scored as Not at all=0, Several days=1, More than half the days=2, and Nearly every day=3; range for score 0 to 27). If more than 1 item is missing, the score should not be calculated. If 1 item is missing, the score is calculated as (sum of answered items*9)/number of answered items.

- **ESRS I (Parkinsonism; Akathisia; Dystonia and Dyskinesia)** = sum of items 1-7 for section I (range for score 0 to 21).

- **ESRS II (Examination: Parkinsonism and Akathisia)** = sum all applicable subsections for section II (range for score 0 to 102).

- **ESRS III (Examination: Dystonia)** = sum of ratings for each of the 10 body parts (range for score 0 to 60).

- **ESRS IV (Examination: Dyskinetic movement)** = sum of 7 items for section IV (range for score 0 to 42).

- **ESRS total score** = sum of each of the 8 sections' scores, where sections V through VIII are individual questions graded on a scale of 0 (absent) to 8 (extremely severe) (range for score 0 to 289).

- **SSS total score** = sum of severity (questions 1 and 2; range for score 3 to 27), locus of control (questions 3, 5, 6; range for score 6 to 54), and avoidance (questions 4, 7, 8; range for score 6 to 54). Only use authority figure and telephone subscores for questions 2-8 (i.e., do not use scores with a close friend) to calculate the total score (range for score 15 to 135).

- **STAI-AD State-Anxiety** = reverse score items 1, 2, 5, 8, 10, 11, 15, 16, 19, and 20, such that 1 → 4, 2 → 3, 3 → 2, 4 → 1 and add all items together for the total score (range for score 20 to 80).

- **STAI-AD Trait-Anxiety** = reverse score items 21, 23, 26, 27, 30, 33, 34, 36, and 39, such that 1 → 4, 2 → 3, 3 → 2, 4 → 1 and add all items together for the total score (range for score 20 to 80).

- **SSI-4 Reading % Syllables Stuttered (SS)** = (# of stuttering events/total number of syllables)*100

- **SSI-4 Reading task score** = assign based on the %SS Reading

  - 1% SS = 2
  - 2% SS = 4
  - 3-4% SS = 5
  - 5-7% SS = 6
  - 8-12% SS = 7
  - 13-20% SS = 8
  - 21+ %SS = 9

- **SSI-4 Speaking % Syllables Stuttered (SS)** = (# of stuttering events/total number of syllables)*100

- **% syllables stuttered** =

  - **For non-readers** = SSI-4 Speaking % SS

  - **For readers** = average of SSI-4 Reading % and Speaking % SS

- **SSI-4 Syllables Stuttered task score** = assigned based on the % SS; for readers, the scale score is determined separately for reading and speaking, and for nonreaders, the scale score is only determined for speaking

  - 1% SS = 2 (readers) or 4 (nonreaders)
  - 2% SS = 3 (readers) or 6 (nonreaders)
  - 3% SS = 4 (readers) or 8 (nonreaders)
  - 4-5% SS = 5 (readers) or 10 (nonreaders)
  - 6-7% SS = 6 (readers) or 12 (nonreaders)
  - 8-11% SS = 7 (readers) or 14 (nonreaders)
  - 13-21% SS = 8 (readers) or 16 (nonreaders)
  - 22+ %SS = 9 (readers) or 18 (nonreaders)

- **SSI-4 frequency scale score** = on a scale of 4 - 18

  - **For readers:** Sum of SS reading task score and SS speaking task score

  - **For nonreaders:** SS speaking task score

- **SSI-4 Duration scale score** = assigned based on the duration score

  - .5 seconds or less = 2
  - .5 - .9 sec = 4
  - 1.0 – 1.9 sec = 6
  - 2.0 – 2.9 sec = 8
  - 3.0 – 4.9 sec = 10

- 5.0 – 9.9 sec = 12
- 10.0 – 29.9 sec = 14
- 30.0 – 59.9 sec = 16
- 60 seconds or more = 18

- **SSI-4 physical concomitants score** = sum of distracting sounds + facial grimaces + head movements + movements of the extremities, on a scale of 0 - 20

- **SSI-4 total score** = sum of frequency scale score + duration scale score + physical concomitants score, on a scale of 6 - 56

### 6.1.8. Data Adjustments/Handling/Conventions

All collected data will be presented in listings. Data not subject to analysis according to this plan will not appear in any tables or graphs but will be included only in the data listings.

All *P* values will be displayed in four decimals and rounded using standard scientific notation (e.g., 0.XXXX). If a *P* value less than 0.0001 occurs, it will be shown in tables as <0.0001; similarly, if a *P* value greater than 0.9999 occurs, it will be shown in tables as >0.9999.

Adverse Events and Concomitant Medication Coding

An AE is the development of an undesirable medical condition or the deterioration of a pre-existing medical condition including worsening of stuttering following or during exposure to a pharmaceutical product, whether or not considered casually related to the product. All AEs that occur after any subject has been screened, enrolled, before treatment, during treatment, or within 30 days following the cessation of treatment, whether or not they are related to the study, will be recorded.

Adverse events will be coded using the MedDRA version 23.0 thesaurus. Conconmitant medications will be coded using the World Health Organization Drug Dictionary (WHO-DD Global B3 version March 2020).

Any event that is likely due to or related to COVID-19, as determined by the investigator, will be marked on the AE CRF as such. These events will be reported separately.

Concomitant Treatment Definition and Handling of Data

A concomitant treatment refers to all treatment, including concomitant therapies as well as herbal treatments, vitamins, behavioral treatment, non-pharmacological treatment, such as psychotherapy, taken between the dates of the first dose of study medication and the end of the subject's participation in the study (Follow-up Day 30), inclusive.

Partial Date Imputation

If partial dates occur, the convention for replacing missing dates for the purpose of statistical analysis is as follows: if just day is missing then the day assigned is the first day of the month or

the date of first dose (if in the same month), whichever is later; if just month is missing then the month assigned is the month of the first dose, unless that results in a date prior to the first dose in which case the month after the first dose is used; and if both month and day are missing then the month assigned is the month of the first dose and the day assigned is either the first day of the month or the first dose date, whichever is later.

If partial times occur, the convention is as follows: if the missing time occurs on the day of the first dose and both the hour and minute are missing then the time assigned is the time of the first dose, otherwise if both the hour and minute are missing and the date is not the date of first dose the time assigned is 12:00; if the date is the same as the date of the first dose and only hour is missing the hour assigned is 12 or the hour of first dose, whichever is later, and if the date is the same as the date of first dose and only the minute is missing the minute assigned is 30 or the minute of first dose, whichever is later. Otherwise the hour assigned is 12 if the hour is missing and the date is not the same as the date of first dose and the minute assigned is 30 is the the date is not the same as the date of first dose.

These conventions will be applied only to adverse event onset dates and times with the following precaution: if the missing date and time reflect the date and time of onset of an adverse event, the modified date and time will be constructed to match the first documented date/time post drug administration while preserving the order in which the AE was reported in the CRF.

Lower and Upper Limit of Quantitation

In general, for quantitative laboratory values reported as "<" or "≤" the lower limit of quantitation (LLOQ), one-half of the reported value (i.e., LLOQ/2) will be used for analysis. The exception to this data treatment is for plasma ecopipam and EBS-101-40853 concentrations that are reported as <LLOQ, where a value of zero will be used in calculating summary statistics.

For quantitative laboratory values reported as ">" or "≥" the upper limit of quantitation (ULOQ), the reported value (i.e., ULOQ) will be used for analysis.

Laboratory Test Results

For analysis purposes, repeat laboratory test results will not be used unless the original laboratory value is missing or indicated as invalid, in which case the first non-missing repeat laboratory value will be used for data analysis.

The International System of Units (SI) will be used in reporting all efficacy and safety laboratory values.

Treatment Duration and Exposure

For subjects who are missing the date of last study drug application, for any reason, the last known contact date will be used in the calculation of treatment duration and study medication exposure. Study drug compliance will not be calculated for those subjects whose date of last study drug application is unknown.

Treatment Phases

TEAE tables will be reported separately for the Titration Phase, Maintenance Phase, and Taper Phase/Follow-Up Period. The Titration Phase will include a 4-week period starting at Baseline (Day 0) through the end of (including) Week 4. The Maintenance Phase will include an 8-week period starting after Week 4 through the end of (including) Week 12. The Completion Phase will include up to one week of a Taper Phase (during which the study drug will be reduced by 25 mg/day until the subject is off the study drug), starting after Week 12 through (including) Follow-Up Day 30.

## 6.2. Special Handling for COVID-19 Disruptions

Study visits may be delayed/not performed as a result of COVID-19 disruptions (e.g., sites were closed or subjects were under stay-at-home orders). Where possible, study sites **may** collect post-baseline data from subjects remotely by telephone or video conference; this will be clearly documented in the source. If possible, sites should adhere to the protocol visit window for remote data collection.

Investigator assessments and subject questionnaires normally completed during on-site visits may be completed via telemedicine or remotely, when applicable due to COVID-19 restrictions The following assessments/questionnaires **may** be collected via telemedicine/remotely only if COVID-19 disruptions or another natural disaster precludes in-person visits:

- SSI-4
- CGI-S
- SSS
- OASES
- C-SSRS
- PHQ-9
- STAI-AD
- Adverse Events
- Conconmitant Medication

The following assessments **cannot** be completed via telemedicine/remotely:

- ESRS
- Urine pregnancy test and drug screen
- PK Blood Draw
- ECG
- Laboratory or urine drug screen tests
- Physical Exam/Vital signs

Study drugs may be delivered to a subject's residence by site personnel who are delegated the responsibility to dispense study drug and coordinate the shipment in the case that restrictions, due to COVID-19 or other natural disaster, prevent in-clinic study visits.

# 7. Study Subjects and Demographics

## 7.1. Disposition of Subjects and Withdrawals

The number of subjects randomized, completed, and withdrawn, along with reasons for withdrawal, will be tabulated overall and by treatment group; the number of subjects screened will also be tabulated by overall only. The number of subjects in each analysis population will be reported.

Study drug accountability and compliance listings will be prepared for all subjects, showing when the planned dosing schedule was not followed, along with the date and type of dosing deviation. Other disposition and study conduct information, including major protocol deviations will be listed.

## 7.2. Protocol Violations and Deviations

Protocol deviations will be listed. Blood samples for PK data will not be considered a protocol deviation if missed at Weeks 4 and 12 due to COVID-19 or other natural disaster situations that may preclude an in-clinic study visit. Certain protocol deviations will be used to exclude subjects from the analyses as discussed in Section 5.

Protocol deviations will be classified as “Important” or “Non-Important”. A major deviation poses a possible safety issue to the subject, or it has a potential impact on the statistical analysis of the clinical data. A minor deviation is identified as any protocol deviation that does not meet the criteria for a major deviation. Major deviations will be reviewed by the Sponsor and Premier to determine the final classification. Protocol deviations which are deemed to be “Important” and “Non-Evaluable” (i.e., a deviation that has a potential impact on the efficacy analysis), will be classified into a separate category. Subjects with at least 1 deviation classified as “Important” and “Non-Evaluable” will be excluded from the analysis population.

Important protocol deviations may include:

- Significant and/or persistent dosing error
- Subject did not meet eligibility criteria and did not have a waiver or dispensation by medical monitor
- Error in randomization (i.e., received wrong drug)
- Use of prohibited concomitant treatment during participation in the trial

Protocol deviations will be summarized by type and by treatment group for the Safety population.

Subjects with a new positive response on questions 4 or 5 of the C-SSRS since last visit version will be immediately discontinued from study drug. Subjects with a PHQ-9 score indicative of the onset of a new depressive episode at any visit will be assessed and may be discontinued from the study drug at the discretion of the Investigator. The Investigator will perform a suicide risk assessment for each of these cases. Dosing may be restarted at the discretion of the Investigator following consultation with the Sponsor or its designee.

The final decision regarding inclusion and exclusion of subjects from the analysis populations will be based on a final listing of protocol deviations. This will be determined during a (blinded) review meeting before any unblinding occurs or database freeze/lock, with input from the Clinical and Biostatistics team members and approval from the Sponsor.

Additionally, inclusion and exclusion criteria not met and reasons for screen failures will be listed.

### 7.3. Demographics and Other Baseline Characteristics

Summary statistics for age, sex (including child-bearing potential), race, ethnicity, height, weight, and weight group, as well as baseline assessments of the CGI-S, OASES overall impact score, OASES IV Quality of Life score, SSS, SSI-4 total score, SSI-4 frequency score, and SSI-4 duration score will be presented by treatment group and overall.

Subjects reporting more than 1 race will be counted in a "More than one race" category for the purposes of tabulating summary statistics for race.

For the continuous variables, the number of non-missing values and the mean, standard deviation, minimum, median and maximum will be tabulated.

For the categorical variables, the counts and proportions of each value will be tabulated.

These analyses will be conducted for the Safety population.

The number and percent of subjects reporting various medical histories, grouped by MedDRA system organ class and preferred term, will be tabulated and presented by treatment group and overall. This analysis will be conducted for the Safety population.

### 7.4. Exposure and Compliance

During the study period, subject compliance will be monitored by review of the tablet counts at the study site visits.

Treatment compliance (%) will be summarized by randomized treatment group and overall. For a given day, a subject is considered compliant with treatment if any amount of study drug was administered. In addition to the presentation of descriptive statistics on compliance rates, treatment compliance will also be categorized and summarized as <80%, 80-120%, and >120%.

Dosing details will be listed separately.

## 8. Efficacy Analysis

Efficacy analyses on the primary, key secondary, and other secondary efficacy endpoints will be performed on the mITT population. The analysis of the primary efficacy endpoint will also be evaluated on the PP population.

For all efficacy analyses, subjects will be analyzed by the randomized treatment group assignment (ecopipam or placebo), based on the intent-to-treat (ITT) principle that asserts that the effect of a treatment can be best assessed by evaluating on the basis of the intention to treat a

subject (i.e., the planned treatment regimen) rather than the actual treatment given. All efficacy data will be presented in subject listings.

## 8.1. Primary Efficacy Analysis

For this study, the primary estimand is the difference in ecopipam and placebo in the mean change from baseline to Week 12 in the SSI-4 total score among all randomized subjects who received at least 1 dose of study drug and have at least 1 post-baseline efficacy evaluation. In the course of the 12-week randomized treatment period, subjects may be exposed to possible known or unknown inter-current events that could possibly impact the estimand, such as treatment discontinuation due to a specific adverse effect or perhaps a lack of effect. The "Treatment Policy Strategy" has been adopted for handling all known or unknown inter-current events in this study. To this end, the ITT principle will serve as the analytical basis for interpreting the estimand. In other words, the difference in ecopipam and placebo in the mean change from baseline to Week 12 in the SSI-4 total score will be evaluated regardless of the occurrence of any such inter-current event.

The primary efficacy endpoint is the change from baseline to Week 12 in the SSI-4 total score. A restricted maximum likelihood (REML)-based MMRM model will be used as the primary analysis method. The repeated measures include post-baseline visits (i.e., Weeks 4, 8, 12) with change from baseline in the SSI-4 total score as the dependent variable. The MMRM model will include the fixed, categorical effects of treatment group, visit, study site, weight-band and treatment group-by-visit interaction as well as the continuous, fixed covariate of baseline SSI-4 total score, using an unstructured covariance structure. If the unstructured covariance model will not converge, the auto-regressive order 1 (AR[1]) and Toeplitz covariance structures will be tested – the structure that converges and has the lowest Akaike information criterion (AIC) will be used in the final model. If the model still does not converge, additional covariance structures will be explored and the final decision will be documented in the CSR.

The primary efficacy analysis will compare ecopipam and placebo using the contrast (difference in least squares [LS] means) between treatment groups through Week 12. Significance tests will be based on the LS means using a 2-sided significance level (2-sided 95% confidence intervals [CIs]).

The null hypothesis for the primary efficacy endpoint of the equality of ecipopam and placebo is:

> $H_{01}$: mean change in total SSI-4 total score between baseline and Week 12 in the two treatment groups are equal.

The null hypothesis of equal treatment effect will be rejected if the statistical analysis results in a 2-sided p-value for treatment over the last 12 weeks of the study of less than or equal to 0.05. Least squares (LS) means will be calculated for each treatment group for each post-baseline visit in the model. The difference between ecopipam and placebo change from baseline in the SSI-4 total score will be estimated, with the corresponding 2-sided 95% CI constructed for each visit. The change from baseline LS means with standard error, 95% CI for the LS means, *P* value for testing if the LS mean is zero, LS mean difference between treatment groups (ecopipam minus placebo) with standard error, 95% CI for the LS mean difference, and *P* value for testing if the treatment LS means are equal will be presented.

The trial will be claimed successful if the hypothesis of no treatment effect on the primary efficacy endpoint of the mITT population is rejected at the 0.05 (2-sided) significance level.

Descriptive summaries will also be provided for the primary endpoint. Data from all study visits will be included in the MMRM analysis, although only results at Week 12 will be included in the fixed-sequence testing strategy to determine study success.

Graphical displays of the calculated change from baseline to Week 12 values in the SSI-4 total scores over time will be presented for each treatment group.

## 8.2. Secondary Efficacy Analyses

Analysis similar to that described for the primary efficacy endpoint will be performed for each of the secondary efficacy endpoints.

The null hypotheses for the secondary efficacy endpoints of the equality of ecopipam and placebo are:

$H_{02}$: mean change in the SSS score between baseline and Week 12 in the two treatment groups are equal

$H_{03}$: mean change in the CGI-S score between baseline and Week 12 in the two treatment groups are equal

$H_{04}$: mean change in the OASES overall impact score between baseline and Week 12 in the two treatment groups are equal

$H_{05}$: proportions of significant improvement in the CGI-I at Week 12 in the two treatment groups are equal

$H_{06}$: mean change in the SSI-4 percentage of syllables stuttered between baseline and Week 12 in the two treatment groups are equal

$H_{07}$: mean change in the SSI-4 frequency subscore score between baseline and Week 12 in the two treatment groups are equal

$H_{08}$: mean change in the SSI-4 duration subscore score between baseline and Week 12 in the two treatment groups are equal

$H_{09}$: mean change in the OASES IV (Quality of Life) sub-score score between baseline and Week 12 in the two treatment groups are equal

$H_{10}$: mean change in the OASES overall impact score sub-score score between baseline and Week 8 in the two treatment groups are equal

$H_{11}$: $H_{02}$: mean change in the SSS score between baseline and Week 8 in the two treatment groups are equal

$H_{12}$: mean change in the CGI-S score between baseline and Week 8 in the two treatment groups are equal

$H_{13}$: mean change in the OASES overall impact score between baseline and Week 8 in the two treatment groups are equal

$H_{14}$: proportions of significant improvement in the CGI-I at Week 8 in the two treatment groups are equal

$H_{15}$: mean change in the SSI-4 percentage of syllables stuttered between baseline and Week 8 in the two treatment groups are equal

$H_{16}$: mean change in the SSI-4 frequency subscore score between baseline and Week 8 in the two treatment groups are equal

$H_{17}$: mean change in the SSI-4 duration subscore score between baseline and Week 8 in the two treatment groups are equal

$H_{18}$: mean change in the OASES IV (Quality of Life) sub-score score between baseline and Week 8 in the two treatment groups are equal

$H_{19}$: mean change in the OASES overall impact score sub-score score between baseline and Week 4 in the two treatment groups are equal

$H_{20}$: $H_{02}$: mean change in the SSS score between baseline and Week 4 in the two treatment groups are equal

$H_{21}$: mean change in the CGI-S score between baseline and Week 4 in the two treatment groups are equal

$H_{22}$: mean change in the OASES overall impact score between baseline and Week 4 in the two treatment groups are equal

$H_{23}$: proportions of significant improvement in the CGI-I at Week 4 in the two treatment groups are equal

$H_{24}$: mean change in the SSI-4 percentage of syllables stuttered between baseline and Week 4 in the two treatment groups are equal

$H_{25}$: mean change in the SSI-4 frequency subscore score between baseline and Week 4 in the two treatment groups are equal

$H_{26}$: mean change in the SSI-4 duration subscore score between baseline and Week 4 in the two treatment groups are equal

$H_{27}$: mean change in the OASES IV (Quality of Life) sub-score score between baseline and Week 4 in the two treatment groups are equal

A hierarchical testing procedure, as described in Section 6.1.3 will be used in the comparisons between ecopipam and placebo on the primary and secondary efficacy endpoints in the mITT population using the primary analysis method (as described for the primary endpoint).

Additionally, the CGI-I and OASES impact scores (overall and IV Quality of Life scores) will be summarized descriptively using frequencies and percentages by study week and treatment group. An analysis will be performed for each visit for the CGI-I in which those with a rating of Very Much Improved or Much Improved (i.e., "Significant Improvement") will be compared to those with all other scores. Proportions will be compared between treatments using Pearson's chi square or Fisher's exact test as appropriate. In addition, odds ratios and corresponding 95% CIs for significant improvement versus all other categories (including missing values at that visit) will be presented. The CGI-S score that is assessed at baseline will be present at each assessment of the CGI-I to ensure the clinician has baseline CGI-S score entered for comparison.

All tests will be performed as 2-sided tests at the 0.05 level of significance.

Graphical displays of the calculated change from baseline to Week 12 values in the key secondary endpoint (CGI-S and OASES overall impact) scores over time will be presented for each treatment group.

### 8.3. Sensitivity and Supportive Analyses for Primary Endpoint

Sensitivity and supportive analyses will be performed on the primary efficacy endpoint to quantify the possible impact of missing data and to demonstrate the robustness of the conclusions. These analyses will only be conducted on the primary efficacy endpoint using the MI model as described above.

Although the assumption of MAR, as used for the primary efficacy analysis method, is often reasonable in clinical trials, the possibility of MNAR data cannot be ruled out. Sensitivity analyses to deal with missing data will use MI methods where missing values are imputed individually under both a plausible MAR and MNAR scenario. A sensitivity analysis based on the MNAR approach (placebo-based imputation) will be carried out to examine the robustness of the MAR assumption in the primary MMRM analysis results (see Section 6.1.4.2).

Summary statistics on the SSI-4 total score by treatment group and visit, including the end of study visit (Week 12), will also be presented.

Subgroups:

Analysis similar to that described for the primary efficacy endpoint will be performed for on the primary endpoint (SSI-4 total score) by several subgroups, including sex, race, SSI-4 total score greater than or equal to/less than the mean score at baseline, and SSI-4 total score greater than or equal to/less than the median score at baseline by study visit and treatment group.

# 9. Safety and Tolerability Analysis

All safety analyses will be performed on the safety population, defined as all subjects who were randomized and received at least 1 dose of study medication.

For all safety and tolerability analyses, subjects will be analyzed by the treatment received (ecopipam or placebo) and, if applicable, overall for all subjects receiving the investigational product (IP). Treatment emergent AEs (TEAEs) will also be summarized by weight band.

Safety data collected at the baseline visit (Visit 2/Day 0) or the last preceding visit if not collected at Visit 2 will be used as the baseline value for safety analyses.

Safety measures including AEs, C-SSRS, PHQ-9, STAI-AD, ESRS, clinical laboratory values, physical examination findings, vital signs, ECGs, weight, and concomitant treatment usage will be summarized descriptively. For quantitative variables, descriptive statistics including number of observations, mean, median, standard deviation, minimum, and maximum will be presented for observed and change from baseline values at each study visit. Qualitative variables will be summarized using counts and percentages.

All safety and tolerability data will be presented in subject listings.

## 9.1. Adverse Events

In general, TEAEs are defined as AEs that start or deteriorate on or after the first dose of study medication and no later than 7 days following the last dose of study medication or reported through the Follow-up Day 30 visit. For any subjects who die during the study and the date of death is between the date of first dose of study medication and the date of study discontinuation (as entered by the site), inclusive, all AEs (including those resulting in death) that occur during the study will be considered as TEAEs irrespective of the last dose and will be included in the TEAE summaries. All summaries of AEs will be based on TEAEs unless specified otherwise.

All AEs, TEAEs, and SAEs will be coded using the Medical Dictionary for Regulatory Activities (MedDRA), Version 23.0 or higher.

Treatment emergent adverse events (TEAEs) are defined as an AE that are newly occurring or worsening after the first dose of study drug. The incidence of TEAEs will be summarized by treatment group, by weight band, by system organ class (SOC) and preferred term (PT), by severity and by relationship to study drug. Serious TEAEs and TEAEs leading to the study termination will also be summarized by SOC, PT, weight band, and treatment group. TEAE safety tables will be reported by Overall TEAEs, Titration Phase TEAEs, Maintenance Phase TEAEs, and Completion Phase TEAEs. The overall summary of Titration Phase, Maintenance Phase and Completion Phase TEAEs will also be presented by each visit. Completion Period TEAEs will include any subsequent Follow Up visits.

The number and percent of subjects reporting TEAEs, grouped by MedDRA SOC and PT (coded using MedDRA v.23.0 or higher), will be tabulated by severity and treatment group. In the case of multiple occurrences of the same TEAE within the same subject, each subject will only be counted once for each preferred term.

In the summaries showing severity and relationship to study medication the event with the maximum severity (severe) or strongest relationship (probably related) will be reported. If a particular event is missing the severity and/or relationship, then the strongest possible severity or relationship will be assumed for analysis (severity = severe, relationship = definitely related).

In the AE data listings, all AEs will be displayed. AEs that are not treatment emergent will be flagged.

### 9.1.1. Adverse Events Leading to Withdrawal

Adverse events leading to study withdrawal will be listed, as well as summarized by system organ class, preferred term, and treatment received (ecopipam or placebo).

### 9.1.2 Adverse Events of Special Interest and Other Adverse Events

Adverse Events of Special Interest (AESI) for ecopipam include standardized MedDRA Queries (SMQs) for MedDRA v. 23.0 (or higher) of Convulsions, Depression and Suicide/Self Injury, Drug Abuse, Dependence and Withdrawal, Dyslipedaemia, EPS, Hepatic Disorders, Hyperglycaemia/New Onset Diabetes Mellitus, Neuroleptic Malignant Syndrome, and Psychosis and Psychotic Disorders.

AESIs for Parkinsonism include Preferred Terms (PTs) of Ataxia, Coordination abnormal, Somnolence, Disturbance in attention, Dyskinesia, Hyperkinesia, Gait disturbance, Memory impairment, Tremor, Vertigo, Agitation, Dystonia, Musculoskeletal stiffness, Dysstasia, Posture abnormal, Balance disorder, Movement disorder, Muscle rigidity, Muscle contractions involuntary, Bradykinesia, Terminal insomnia, Nightmare, Poor quality sleep, Sleep disorder, Fatigue, Dizziness, Restlessness, Amnesia, Confusional state, Dementia, Mental impairment, Speech disorder, Dysphonia, Vocal cord dysfunction, Parosmia, Anosmia, Urinary incontinence, Anxiety, Apathy, Joint stiffness, Reduced facial expression, Flat affect, Constipation, Dysphagia, Drooling, Fall, Fear of falling, Loss of visual contrast sensitivity, Muscle tightness, Dysgraphia, and Weight decreased.

AESI summary tables are listed below:

1) AESIs for ecopipam overall and by system organ class and preferred term

2) AESIs for Parkinsonism overall and by system organ class and preferred term.

Significant AEs of particular clinical importance, other than SAEs and those AEs leading to discontinuation of the subject/subjects from the study, will be classified as Other Adverse Events (OAEs).

### 9.1.3 Deaths and Serious Adverse Events

Any deaths that occur during the study will be listed.

Serious adverse events (SAEs) will be listed and will also tabulated by system organ class and preferred term and presented by treatment group, severity, and relationship to study drug.

An SAE is an AE occurring during any study phase (i.e., baseline, treatment, washout, or follow-up), and at any dose of the investigational product, comparator or placebo, that fulfills one or more of the following:

- Results in death
- It is immediately life-threatening
- It requires in-subject hospitalization or prolongation of existing hospitalization
- It results in persistent or significant disability or incapacity
- Results in a congenital abnormality or birth defect
- It is an important medical event that may jeopardize the subject or may require medical intervention to prevent one of the outcomes listed above

## 9.2 Clinical Laboratory Evaluations

Central laboratory tests (hematology, chemistry, electrolytes, liver function tests, renal function parameters, HbA1c, urine) will be performed at Screening, Baseline, Week 4, Week 8, Week 12 (or early termination), and Follow-up Day 7 and Day 14, with the exception of HbA1c, which will be performed at Baseline and Week 12 only. Laboratory test results will be summarized descriptively by treatment and visit as both observed values and change from Baseline values.

The number of subjects with clinical laboratory values below, within, or above the normal range by study visit and in relation to baseline will be tabulated for clinical chemistry and HbA1c by treatment group (shift table).

Local laboratory tests (urine pregnancy test and urine drug screen) will be performed at Screening, Baseline, and Week 12 (or early termination). These results will not be reported unless applicable for enrollment/discontinuation reasons.

## 9.3 Vital Signs

Vital signs will be collected at Screening, Baseline, Week 4, Week 8, Week 12 (or early termination), and Follow-up Day 7 and Day 14. Descriptive summaries of actual values and changes from baseline will be calculated for supine pulse (bpm), supine systolic blood pressure (mmHg), supine diastolic blood pressure (mmHg), standing pulse (bpm), standing systolic blood pressure (mmHg), standing diastolic blood pressure (mmHg), weight (kg), and height (cm). Height is measured at Screening only.

## 9.4 Electrocardiograms

12-Lead ECGs will be collected at Screening, Baseline, Week 4, Week 8, Week 12, Follow-up Day 7, and Follow-up Day 14. Descriptive summaries will be presented for heart rate (bpm), PR interval (ms), RR interval (ms), QRS (ms), uncorrected QT interval (ms), QTcB (ms), QTcF (ms). Descriptive summaries will be presented by study visit and treatment group.

The number and percentage of subjects with normal, abnormal but not clinically significant, and abnormal and clinically significant ECG results will be summarized by treatment group and study visit.

## 9.5 Physical Examinations

The number and percentage of subjects with normal and abnormal findings in the physical examination at baseline and each post-baseline visit will be displayed for each treatment group. A listing of physical examination findings will be provided for complete and brief examiniatons.

## 9.6 Further Safety Evaluations

### 9.6.2 Columbia Suicide Severity Rating Scale (C-SSRS)

The C-SSRS is an instrument to assess both suicidal behavior and ideation. The Screening/Baseline version will be administered at Screening and the Since Last Visit version of the scale will be administered at all subsequent visits (Week 2, Week 4, Week 6, Week 8, Week 10, Week 12, and Follow-Up Days 7 and 14). Results will be summarized by treatment group and study visit.

### 9.6.3 Patient Health Questionnaire-9 (PHQ-9)

The PHQ-9 is a validated, 9-question tool to assess for the degree of depression present in an individual and will be administered at Screening, Baseline, Week 4, Week 8, and Week 12. The PHQ-Score will be summarized by treatment group and study visit.

### 9.6.4 State-Trait Anxiety Index-Adult (STAI-AD)

The STAI-AD is a commonly used self-report to measure two types of anxiety: state anxiety (or anxiety about an event), and trait anxiety (or anxiety level as a personal characteristic). The STAI-AD has 40 items, 20 items allocated to each of the State-Anxiety and Trait-Anxiety subscales. The STAI-AD is a psychological inventory based on a 4-point Likert scale. Responses for the S-Anxiety scale assess intensity of current feelings “at this moment”: 1) not at all, 2) somewhat, 3) moderately so, and 4) very much so. Responses for the T-Anxiety scale assess frequency of feelings “in general”: 1) almost never, 2) sometimes, 3) often, and 4) almost always. Higher scores are positively correlated with higher levels of anxiety.

The Trait-Anxiety score at Screening and the State-Anxiety score at Baseline, Week 4, Week 8, and Week 12 will be summarized by treatment group and study visit.

### 9.6.5 Extra-Pyramidal Symptom Rating Scale (ESRS)

The ESRS is a clinician-reported instrument developed to assess four types of drug-induced movement disorders (DIMD): Parkinsonism, akathisia, dystonia, and tardive dyskinesia (TD). The ESRS consists of:

14) four subscales:

- a questionnaire of EPS or DIMD
- an examination of Parkinsonism and akathisia
- an examination of dystonia
- an examination of dyskinesia

B) four CGI-S scales assessing:

- TD
- Parkinsonism
- Dystonia
- akathisia

A score is calculated for each of the 8 sub-categories by summing the scores of the individual questions, and these scores are summarized by treatement group and study visit (Baseline, Week 4, Week 8, and Week 12).

### 9.7 Concomitant Medications

Prior and concomitant treatments will be summarized descriptively by treatment group using the number and proportion of subjects by preferred term and ATC Class Level 4.

Prior treatments will be summarized separately from concomitant treatments. Treatments that started before the first dose of study medication will be considered prior treatments, whether or not they were stopped before the first dose of study medication. Any treatments continuing or starting after the first dose of study medication will be considered to be concomitant. If a treatment starts before the first dose of study medication and continues after the first dose of study medication, it will be considered both prior and concomitant.

Prior and concomitant procedures will be listed separately from medications; these were coded using MedDRA.

## 10 Changes from Planned Analysis

Summary tables for frequencies of OASES Overall Impact and section IV Quality of Life Impact scores will be included.

## 11 Population Pharmacokinetic (PK) Analysis

Systemic concentrations of ecopipam and its major (active) metabolite, EBS-101-40853 in plasma will be determined at pre-dose (20-36 hours since the last dose), post-dose between 0.5 and 1.5 hours, and post-dose between 2 and 4 hours at Week 4. Any samples will be collected at least 30 minutes apart. A single sample will be collected at Week 12. Any PK blood samples not collected due to restrictions from COVID-19 or another natural disaster will not be considered a protocol deviation. Any data collected on the CRF will be provided in a listing.

A data dependent population PK analysis will be performed under a separate Bioanalytical plan.

## 12 References

1. US Federal Register. (1998) International Conference on Harmonization; Guidance on Statistical Principles for Clinical Trials. Department of Health and Human Services: Food and Drug Administration [Docket No. 97D-0174]. Federal Register Volume 63, Number 179, pages 49583-49598. September 16, 1998.

2. ASA. (2016) Ethical Guidelines for Statistical Practice. Prepared by the Committee on Professional Ethics, April 2016. http://www.amstat.org/about/ethicalguidelines.cfm

3. RSS. (2014) The Royal Statistical Society: Code of Conduct, 2014. http://www.rss.org.uk/Images/PDF/join-us/RSS-Code-of-Conduct-2014.pdf.

4. Mallinckrodt C. H., Lane P. W., Schnell D., Peng Y., Mancuso J. P. 2008. Recommendations for the Primary Analysis of Continuous Endpoints in Longitudinal Clinical Trials. *Drug Information Journal*, 42, 303-319.

5. Molenberghs G., Kenward M. G. 2007. *Missing Data in Clinical Studies.* New York: Wiley.

6. Cnaan A., Laird N. M., Slasor P. 1997. Using the general linear mixed model to analyse unbalanced repeated measures and longitudinal data. *Stat Med.*, 16, 2349-2380.

7. Molenberghs G., Thijs H., Jansen I., et al. 2004. Analyzing incomplete longitudinal clinical trial data. *Biostatistics*, 5, 445-464.

8. Mallinckrodt C. H., Clark W. S., David S. R. 2001. Accounting for dropout bias using mixed-effects models. *Journal of Biopharmaceuticals Statistics,* 11, 9-21.

9. European Medicines Agency. 2010. Guideline on Missing Data in Confirmatory Clinical Trials.

10. Siddiqui O., Hung H. M., O'Neill R. 2009. MMRM vs. LOCF: a comprehensive comparison based on simulation study and 25 NDA datasets. *Journal of Biopharmaceuticals Statistics*, 19(2), 227-246.

11. Mallinckrodt C. H., Roger J., Chuang-Stein C., Molenberghs G., O'Kelly M., Ratitch B., Janssens M., Bunouf P. 2013. Recent Developments in the Prevention and Treatment of Missing Data. *Drug Information Journal*, 48, 68-80.

12. Little R., Rubin D. 1987. *Statistical Analysis with Missing Data*. New York: John Wiley.

13. Verbeke G., Molenberghs G. 2000. *Linear Mixed Models for Longitudinal Data*. New York: Springer.

14. Ratitich, B. and M. O'Kelly (2011). "Implementation of Pattern-Mixture Models Using Standard SAS/STAT Procedures." PharmaSUG. Available at http://www.pharmasug.org/proceedings/2011/SP/PharmaSUG-2011-SP04.pdf

# 13 Tables, Listings, and Figures

All listings, tables, and graphs will have a header showing the sponsor company name (Emalex Biosciences, Inc.) and protocol and a footer showing the version of SAS, the file name and path, and the source of the data (listing or table number, as applicable).

The following reporting conventions will be adopted for the presentation of study data. These conventions will enhance the review process and help to standardize the presentation with common notations.

**General Reporting Conventions**

- All tables and data listings will be developed in landscape orientation.
- Specialized text styles, such as bolding, italics, borders, shading, and superscripted and subscripted text will not be used in tables, figures, and data listings unless they add significant value to the table, figure, or data listing.
- Only standard keyboard characters should be used in tables and data listings. Special characters, such as nonprintable control characters, printer-specific characters, or font specific characters, will not be used on a table, figure, or data listing.
- Adverse events with missing MedDRA coding will have their system organ class and/or preferred term presented as "Not Coded" in the tables. The "Not Coded" frequencies will be sorted to the end of the tables. Adverse events that are not coded are not expected.
- Programming notes may be inserted into the shells, these notes will not appear in the final output.

**Population Summary Conventions**

- Population sizes may be presented for each classification factor as totals in the column header as (N=xx), where appropriate.
- All population summaries for categorical variables will include all categories that were planned and for which the subjects may have had a response. Percentages corresponding to null categories (cells) will be suppressed; however, counts and percentages of missing values may be needed. Counts of zero will be presented as '0', without the percentage.
- All population summaries for continuous variables will include: n, mean, SD, median, minimum, and maximum. Other summaries (e.g., number missing, geometric mean, median, quartiles, 95% CIs, and coefficient of variation [CV] or % CV) may be used as appropriate. The precision of the maximum and minimum will match the maximum precision in the data. The mean and median will have 1 additional decimal place. The SD will have 2 additional decimal places.
- All percentages are rounded and reported to a single decimal point (xx.x%).

## 13.1 Planned Table Descriptions

The following are planned summary tables for protocol number EBS-101-COFD-201. The table numbers and page numbers are place holders only and will be determined when the tables are

produced. The efficacy outputs are ordered to account for the hierarchical testing structure as outlined in Section 6.1.3, where possible. For cases where multiple study visits for the same parameter are tested, those visits are consolidated on a single table for ease of review and referencing results. Similarly, for cases where multiple analysis populations or cohort groups are presented, this data will be consolidated on a single table, when practical.

**Table 4: Demographic Data Summary Tables and Figures**

| Table Number | Population | Title |
|---|---|---|
| **14.1 Demographic Data Summary Tables and Figures** | | |
| 14.1.1 | All | Subject Disposition |
| 14.1.2 | Safety | Demographics and Baseline Characteristics |
| 14.1.3 | Safety | Summary of Medical and Psychiatric History |
| 14.1.4 | Safety | Summary of Prior Medication |
| 14.1.5 | Safety | Study Drug Exposure by Phase |

**Table 5: Efficacy Data**

| Table Number | Population | Title |
|---|---|---|
| **14.2 Efficacy Tables** | | |
| 14.2.1.1 | mITT | Mean Change from Baseline in SSI-4 Total Score by Study Visit and Treatment – Mixed Model for Repeated Measures Analysis – Primary Analysis |
| 14.2.1.2 | PP | Mean Change from Baseline in SSI-4 Total Score by Study Visit and Treatment – Mixed Model for Repeated Measures Analysis – Primary Analysis |
| 14.2.1.3 | mITT | Mean Change from Baseline in SSI-4 Total Score by Study Visit and Treatment – Sensitivity Analysis using MI under MNAR |
| 14.2.1.4 | mITT | Mean Change from Baseline in SSI-4 Subscales by Study Visit and Treatment – Mixed Model for Repeated Measures Analysis |
| 14.2.1.5 | mITT | Summary of SSI-4 Total Score by Study Visit, Treatment, and Sex – Mixed Model for Repeated Measures Analysis – Subgroup Analysis of the Primary Endpoint |
| 14.2.1.6 | mITT | Summary of SSI-4 Total Score by Study Visit, Treatment, and Race – Mixed Model for Repeated Measures Analysis – Subgroup Analysis of the Primary Endpoint |
| 14.2.1.7 | mITT | Summary of SSI-4 Total Score by Study Visit, Treatment, and Mean Split – Mixed Model for Repeated Measures Analysis – Subgroup Analysis of the Primary Endpoint |
| 14.2.1.8 | mITT | Summary of SSI-4 Total Score by Study Visit, Treatment, and Median Split – Mixed Model for Repeated Measures Analysis – Subgroup Analysis of the Primary Endpoint |
| 14.2.2 | mITT | Mean Change from Baseline in SSS by Study Visit and Treatment – Mixed Model for Repeated Measures Analysis |
| 14.2.3 | mITT | Mean Change from Baseline in CGI-S by Study Visit and Treatment – Mixed Model for Repeated Measures Analysis |
| 14.2.4.1 | mITT | Mean Change from Baseline in OASES Overall Impact score by Study Visit and Treatment – Mixed Model for Repeated Measures Analysis |

| Table Number | Population | Title |
|---|---|---|
| 14.2.4.2 | mITT | Summary of OASES Overall Impact Score by Study Visit and Treatment |
| 14.2.5.1 | mITT | Summary of Clinical Global Impression – Improvement (CGI-I) by Study Visit and Treatment |
| 14.2.5.2 | mITT | Analysis of Clinical Global Impression – Improvement (CGI-I) by Study Visit – Significant Improvement vs All Other Categories |
| 14.2.6.1 | mITT | Mean Change from Baseline in OASES Section IV Quality of Life score by Study Visit and Treatment – Mixed Model for Repeated Measures Analysis |
| 14.2.6.2 | mITT | Summary of OASES Section IV Quality of Life Impact Score by Study Visit and Treatment |

## Table 6: Safety Data

| Number | Population | Title |
|---|---|---|
| **14.3.1 Displays of Adverse Events** | | |
| 14.3.1.1.1 | Safety | Summary of Treatment Emergent Adverse Events – Overall |
| 14.3.1.1.2 | Safety | Summary of Treatment Emergent Adverse Events – Titration Phase |
| 14.3.1.1.3 | Safety | Summary of Treatment Emergent Adverse Events – Maintenance Phase |
| 14.3.1.1.4 | Safety | Summary of Treatment Emergent Adverse Events – Completion Phase |
| 14.3.1.1.5 | Safety | Summary of Treatment Emergent Adverse Events by Study Visit – Titration Phase |
| 14.3.1.1.6 | Safety | Summary of Treatment Emergent Adverse Events by Study Visit – Maintenance Phase |
| 14.3.1.1.7 | Safety | Summary of Treatment Emergent Adverse Events by Study Visit – Completion Phase |
| 14.3.1.2.1 | Safety | Incidence of Treatment Emergent Adverse Events by System Organ Class and Preferred Term – Overall |
| 14.3.1.2.2 | Safety | Incidence of Treatment Emergent Adverse Events by System Organ Class and Preferred Term – Titration Phase |
| 14.3.1.2.3 | Safety | Incidence of Treatment Emergent Adverse Events by System Organ Class and Preferred Term – Maintenance Phase |
| 14.3.1.2.4 | Safety | Incidence of Treatment Emergent Adverse Events by System Organ Class and Preferred Term – Completion Phase |
| 14.3.1.3.1 | Safety | Incidence of Treatment Emergent Adverse Events by System Organ Class, Preferred Term, and Maximum Severity – Overall |
| 14.3.1.3.2 | Safety | Incidence of Treatment Emergent Adverse Events by System Organ Class, Preferred Term, and Maximum Severity – Titration Phase |
| 14.3.1.3.3 | Safety | Incidence of Treatment Emergent Adverse Events by System Organ Class, Preferred Term, and Maximum Severity – Maintenance Phase |
| 14.3.1.3.4 | Safety | Incidence of Treatment Emergent Adverse Events by System Organ Class, Preferred Term, and Maximum Severity – Completion Phase |

| Number | Population | Title |
|---|---|---|
| 14.3.1.4.1 | Safety | Incidence of Treatment Emergent Adverse Events by System Organ Class, Preferred Term, and Relationship to Study Drug – Overall |
| 14.3.1.4.2 | Safety | Incidence of Treatment Emergent Adverse Events by System Organ Class, Preferred Term, and Relationship to Study Drug – Titration Phase |
| 14.3.1.4.3 | Safety | Incidence of Treatment Emergent Adverse Events by System Organ Class, Preferred Term, and Relationship to Study Drug – Maintenance Phase |
| 14.3.1.4.4 | Safety | Incidence of Treatment Emergent Adverse Events by System Organ Class, Preferred Term, and Relationship to Study Drug –Completion Phase |
| **14.3.2 Summary of Deaths, Other Serious and Significant Adverse Events, and AEs of Special Interest** | | |
| Table 14.3.2.1.1 | Safety | Incidence of Serious Adverse Events by System Organ Class and Preferred Term – Overall |
| Table 14.3.2.1.2 | Safety | Incidence of Serious Adverse Events by System Organ Class and Preferred Term – Titration Phase |
| Table 14.3.2.1.3 | Safety | Incidence of Serious Adverse Events by System Organ Class and Preferred Term – Maintenance Phase |
| Table 14.3.2.1.4 | Safety | Incidence of Serious Adverse Events by System Organ Class and Preferred Term – Completion Phase |
| Table 14.3.2.2.1 | Safety | Incidence of Serious Adverse Events by System Organ Class, Preferred Term, and Relationship to Study Drug – Overall |
| Table 14.3.2.2.2 | Safety | Incidence of Serious Adverse Events by System Organ Class, Preferred Term, and Relationship to Study Drug – Titration Phase |
| Table 14.3.2.2.3 | Safety | Incidence of Serious Adverse Events by System Organ Class, Preferred Term, and Relationship to Study Drug – Maintenance Phase |
| Table 14.3.2.2.4 | Safety | Incidence of Serious Adverse Events by System Organ Class, Preferred Term, and Relationship to Study Drug – Completion Phase |
| Table 14.3.2.3.1 | Safety | Incidence of Treatment Emergent Adverse Events of Special Interest – Ecopipam-Overall |
| Table 14.3.2.3.2 | Safety | Incidence of Treatment Emergent Adverse Events of Special Interest – Ecopipam–Titration Phase |
| Table 14.3.2.3.3 | Safety | Incidence of Treatment Emergent Adverse Events of Special Interest – Ecopipam–Maintenance Phase |
| Table 14.3.2.3.4 | Safety | Incidence of Treatment Emergent Adverse Events of Special Interest – Ecopipam–Completion Phase |
| Table 14.3.2.4.1 | Safety | Incidence of Treatment Emergent Adverse Events of Special Interest – Parkinsonism – Overall |
| Table 14.3.2.4.2 | Safety | Incidence of Treatment Emergent Adverse Events of Special Interest – Parkinsonism – Titration Phase |
| Table 14.3.2.4.3 | Safety | Incidence of Treatment Emergent Adverse Events of Special Interest – Parkinsonism – Maintenance Phase |

| Number | Population | Title |
|---|---|---|
| Table 14.3.2.4.4 | Safety | Incidence of Treatment Emergent Adverse Events of Special Interest – Parkinsonism – Completion Phase |
| **14.3.3 Narratives of Deaths, Other Serious and Certain Other Significant Adverse Events** | | |
| Table 14.3.3.1 | Safety | Listing of Adverse Events Leading to Study Drug Discontinuation |
| Table 14.3.3.2 | Safety | Listing of Serious Adverse Events |
| Table 14.3.3.3 | Safety | Listing of Deaths |
| **14.3.4 Abnormal Laboratory Values** | | |
| Table 14.3.4.1 | Safety | Listing of Abnormal Laboratory Data |
| **14.3.5 Laboratory Data Summary Tables** | | |
| 14.3.5.1 | Safety | Laboratory Data: Change from Baseline in Hematology Parameters by Study Visit |
| 14.3.5.2 | Safety | Laboratory Data: Change from Baseline in Chemistry Parameters by Study Visit |
| 14.3.5.2.1 | Safety | Laboratory Data: Shift from Baseline in Chemistry Parameters by Study Visit |
| 14.3.5.3 | Safety | Laboratory Data: Change from Baseline in Electrolyte Parameters by Study Visit |
| 14.3.5.4 | Safety | Laboratory Data: Change from Baseline in Liver and Renal Function Parameters by Study Visit |
| 14.3.5.5 | Safety | Laboratory Data: Change from Baseline in Urine Parameters by Study Visit |
| 14.3.5.6 | Safety | Laboratory Data: Frequency of Categorical Urinalysis Parameters by Study Visit |
| 14.3.5.7 | Safety | Laboratory Data: Change from Baseline in HbA1c by Study Visit |
| 14.3.5.7.1 | Safety | Laboratory Data: Shift from Baseline in HbA1c by Study Visit |
| **14.3.6 Other Safety Data Summary Tables** | | |
| 14.3.6.1 | Safety | Summary of Vital Signs by Study Visit |
| 14.3.6.2 | Safety | Summary of 12-Lead Electrocardiogram by Study Visit |
| 14.3.6.2.1 | Safety | Summary of 12-Lead Electrocardiogram Interpretation by Study Visit |
| 14.3.6.3.1 | Safety | Summary of Suicidal Ideation, Suicidal Behavior, and Suicide based on C-SSRS |
| 14.3.6.4 | Safety | Summary of Extra-Pyramidal Symptom Rating Scale Questionnaire |
| 14.3.6.5 | Safety | Summary of Patient Health Questionnaire-9 score |
| 14.3.6.6 | Safety | Summary of State-Trait Anxiety Index-Adult Score |
| 14.3.6.7 | Safety | Summary of Physical Examination by Study Visit |
| 14.3.6.8 | Safety | Summary of Concomitant Medications by ATC Class Level 4 and Preferred Term |

## 13.2 Planned Listing Descriptions

The following are planned data and subject data listings for protocol number EBS-101-COFD-201.

In general, one listing will be produced per CRF domain. All listings will be sorted by treatment, site, and subject number. All calculated variables will be included in the listings.

In all listings a blank line will be placed between each subject. Within a data listing, if an item appears line after line (e.g., repetition of subject number), then only the first occurrence will be displayed.

In data listings, the information for one subject will be kept on one page if at all possible, rather than splitting a subject's information across pages.

**Table 7: Planned Listings**

| Number | Dataset | Title / Summary |
|---|---|---|
| **16.2.1 Subject Discontinuations/Completions** | | |
| Listing 16.2.1 | All Subjects | Subject Disposition |
| **16.2.2 Protocol Deviations** | | |
| Listing 16.2.2.1 | All Subjects | Inclusion and Exclusion Criteria and Re-Consent |
| Listing 16.2.2.2 | All Subjects | Protocol Deviations |
| Listing 16.2.2.3 | Screen Failures | Resasons for Screen Failures |
| **16.2.3 Subjects Excluded from the Efficacy Analyses** | | |
| Listing 16.2.3 | All Subjects | Analysis Populations |
| **16.2.4 Demographic Data and Other Baseline Characteristics** | | |
| Listing 16.2.4.1 | All Subjects | Subject Consent and Demographics |
| Listing 16.2.4.2 | All Subjects | Medical History |
| Listing 16.2.4.3 | All Subjects | Psychiatric History |
| Listing 16.2.4.4 | All Subjects | Complete Physical Examination |
| Listing 16.2.4.5 | All Subjects | Brief Physical Examination |
| **16.2.5 Compliance and/or Drug Concentration Data** | | |
| Listing 16.2.5.1 | All Subjects | Titration/Reserve Kit Dispensation and Accountability |
| Listing 16.2.5.2 | All Subjects | Maintenance Dispensation and Accountability |
| Listing 16.2.5.3 | All Subjects | Taper Phase Dispensation and Accountability |
| Listing 16.2.5.4 | All Subjects | PK Assessment at Week 4 |
| Listing 16.2.5.5 | All Subjects | PK Assessment at Week 12 |
| **16.2.6 Individual Efficacy Response Data** | | |
| Listing 16.2.6.1 | All Subjects | Stuttering Severity Instrument – 4 (SSI-4) |

| Number | Dataset | Title / Summary |
|---|---|---|
| Listing 16.2.6.2 | All Subjects | Clinical Global Impression – Severity (CGI-S) |
| Listing 16.2.6.3 | All Subjects | Clinical Global Impression – Improvement (CGI-I) |
| Listing 16.2.6.4 | All Subjects | Overall Assessment of the Speaker's Experience of Stuttering (OASES) Scores |
| Listing 16.2.6.5 | All Subjects | Subjective Screening of Stuttering (SSS) |
| **16.2.7 Adverse Event Listings (by Subject)** | | |
| Listing 16.2.7.1 | All Subjects | Adverse Events |
| **16.2.8 Laboratory Values (by Subject)** | | |
| Listing 16.2.8.1 | All Subjects | Laboratory Data: Hematology |
| Listing 16.2.8.2 | All Subjects | Laboratory Data: Chemistry |
| Listing 16.2.8.3 | All Subjects | Laboratory Data: Urinalysis |
| Listing 16.2.8.4 | All Subjects | Urine Drug Screen |
| Listing 16.2.8.5 | Female Subjects | Urine Pregnancy Test |
| **16.2.9 Other Clinical Observations and Measurements (by Subject)** | | |
| Listing 16.2.9.1.1 | All Subjects | Prior and Concomitant Medications |
| Listing 16.2.9.1.2 | All Subjects | Prior and Concomitant Procedures |
| Listing 16.2.9.2 | All Subjects | Vital Signs |
| Listing 16.2.9.3 | All Subjects | ECG Measurements |
| Listing 16.2.9.4 | All Subjects | COVID-19 Impact Log |
| Listing 16.2.9.5 | All Subjects | COVID-19 IP Impact |
| **16.2.10 Other Study Measurements or Assessments (by Subject)** | | |
| Listing 16.2.10.1 | All Subjects | Diagnostic and Statistical Manual for Mental Disorders, 5th Edition (DSM-5) |
| Listing 16.2.10.2 | All Subjects | Columbia Suicide Severity Rating Scale (C-SSRS) Screening |
| Listing 16.2.10.3 | All Subjects | Columbia Suicide Severity Rating Scale (C-SSRS) Since Last Visit |
| Listing 16.2.10.4 | All Subjects | Extrapyramidal Symptom Rating Scale (ESRS) |
| Listing 16.2.10.5 | All Subjects | Patient Health Questionnaire-9 (PHQ-9) |
| Listing 16.2.10.6 | All Subjects | State-Trait Anxiety Inventory for Adults (STAI-State) |
| Listing 16.2.10.7 | All Subjects | State-Trait Anxiety Inventory for Adults (STAI-Trait) at Baseline Visit |

## 13.3 Planned Figure Descriptions

The following are planned summary figures for protocol number EBS-101-COFD-201. The figure numbers and page numbers are place holders only and will be determined when the figures are produced.

**Table 8: Planned Figures**

| Figure Number | Population | Figure Title/Summary |
|---|---|---|
| 14.2.1.1 | mITT | Plot of Mean Change from Baseline in SSI-4 Total Score |
| 14.2.1.2 | mITT | Plot of SSI-4 Total Score Tipping Point Analysis P Values vs. Shift Parameters |
| 14.2.2 | mITT | Plot of Mean Change from Baseline in CGI-S |
| 14.2.3 | mITT | Plot of Mean Change from Baseline in SSS Score |
| 14.2.4 | mITT | Plot of Mean Change from Baseline in OASES Overall Impact Score |
| 14.2.5 | mITT | Plot of Summary of CGI-I |
| 14.3.1 | mITT | Plot of Summary of PHQ-9 Scores |
| 14.3.2 | mITT | Plot of Summary of STAI-State Scores |