# STATISTICAL ANALYSIS PLAN

## SEP361-203

**A MULTICENTER RANDOMIZED DOUBLE-BLIND FOLLOWED BY AN OPEN-LABEL EXTENSION STUDY TO EVALUATE THE EFFICACY, SAFETY, AND TOLERABILITY OF SEP-363856 IN SUBJECTS WITH PARKINSON'S DISEASE PSYCHOSIS**

## AUTHOR:

## VERSION NUMBER AND DATE: FINAL V1, 24FEB2020

## STATISTICAL ANALYSIS PLAN SIGNATURE PAGE

Statistical Analysis Plan Final v1.0 (Dated 24FEB2020) for Protocol SEP361-203.

| | Name | Signature | Date |
|---|---|---|---|
| **Author:** | | | |
| **Position:** | | | |
| **Company:** | | | |
| | | | |
| **Approved By:** | | | |
| **Position:** | | | |
| **Position:** | | | |

Upon review of this document, the undersigned approves this version of the Statistical Analysis Plan, authorizing that the content is acceptable for the reporting of this study.

| | Name | Signature | Date |
|---|---|---|---|
| **Approved By:** | | | |
| **Position:** | | | |
| **Company:** | | | |
| | | | |
| **Approved By:** | | | |
| **Position:** | | | |
| **Company:** | | | |
| | | | |
| **Approved By:** | | | |
| **Position:** | | | |
| **Company:** | | | |

| | | | |
|---|---|---|---|
| | | | |
| **Approved By:** | | | |
| **Position:** | | | |
| **Company:** | | | |
| | | | |
| **Approved By:** | | | |
| **Position:** | | | |
| **Company:** | | | |

| | | | |
|---|---|---|---|
| **Approved By:** | | | |
| **Position:** | | | |
| **Company:** | | | |
| | | | |
| **Approved By:** | | | |
| **Position:** | | | |
| **Company:** | | | |

## MODIFICATION HISTORY

| Unique Identifier for this Version | Date of the Document Version | Author | Significant Changes from Previous Authorized Version |
|---|---|---|---|
| Draft 1.0 | 16NOV2017 | | Not Applicable – First Version |
| Draft 2.0 | 20MAR2018 | | Update as per sponsor comments on version 1.0 |
| Draft 3.0 | 31AUG2019 | | Update as per protocol V6.0 |
| Draft 3.2 | 20DEC2019 | | Implement Sunovion comments |
| Final 1.0 | 24FEB2020 | | Implement Sunovion comments |

# TABLE OF CONTENTS

## LIST OF ABBREVIATIONS

| Abbreviation | Explanation |
|---|---|
| **AE** | Adverse Event |
| **ANCOVA** | Analysis of Covariance |
| **ATC** | Anatomical Therapeutic Chemical |
| **BMI** | Body Mass Index |
| **BLQ** | Below the Lower Limit of Quantification |
| **BPST-PD** | Brief Psychosocial Therapy adapted for Parkinson's Diseases |
| **C-SSRS** | Columbia Suicide Severity Rating Scale |
| **CGI-S** | Clinical Global Impression – Severity of Illness |
| **CI** | Confidence Interval |
| **CRF** | Case Report Form |
| **CS** | Clinically Significant |
| **CSP** | Clinical Study Protocol |
| **DMC** | Data Monitoring Committee |
| **ECG** | Electrocardiogram |
| **ENR** | All Subjects Enrolled |
| **HLT** | High Level Term |
| **HR** | Heart Rate |
| **IPD** | Important Protocol Deviations |
| **IXRS** | Interactive voice/web response system |
| **LOCF** | Last Observation Carried Forward |
| **LS** | Least Squares |
| **MAR** | Missing At Random |
| **MCMC** | Markov Chain Monte Carlo |
| **MedDRA** | Medical Dictionary for Regulatory Activities |

| Abbreviation | Explanation |
|---|---|
| **mITT** | Modified Intent-to-Treat |
| **mITTComp** | Modified Intent-to-Treat Completers |
| **MMRM** | Mixed Model for Repeated Measures |
| **MMSE** | Mini mental state examination |
| **NCS** | Not Clinically Significant |
| **NE** | Neurological Examination |
| **NNT** | Number Needed to Treat |
| **NPI** | Neuropsychiatric Inventory |
| **NPI (H+D)** | Neuropsychiatric Inventory (Hallucination + Delusion) |
| **PCS** | Potentially Clinically Significant |
| **PDP** | Parkinson's disease psychosis |
| **PMM** | Pattern Mixture Model |
| **PD** | Pharmacodynamic |
| **PE** | Physical Examination |
| **PK** | Pharmacokinetic |
| **PP** | Per Protocol |
| **PT** | Preferred Term |
| **QD** | one a day (from the Latin quaque die) |
| **REM** | Rapid Eye Movement |
| **RBD** | Sleep Behavioral Disorder |
| **RND** | All Subjects Randomized |
| **RR** | Risk Reduction |
| **SAE** | Serious Adverse Event |
| **SAF** | Safety Population |
| **SAP** | Statistical Analysis Plan |

| Abbreviation | Explanation |
|---|---|
| **SAPS-PD** | Scale for Assessment of Positive Symptoms – Parkinson's Disease |
| **SCOPA-DS** | Scales for Outcomes in Parkinson's Disease for daytime sleepiness |
| **SCOPA-NS** | Scales for Outcomes in Parkinson's Disease for nighttime sleep |
| **SCR** | Screened Population |
| **SE** | Standard Error |
| **SI** | International System of Units |
| **SOC** | System Organ Class |
| **ULQ** | Above the Upper Limit of Quantification |
| **UPDRS** | Unified Parkinson's Disease Rating scale |
| **WHO** | World Health Organization |
| **vs** | versus |

# 1. INTRODUCTION

This document describes the rules and conventions to be used in the presentation and analysis of efficacy and safety data for Protocol SEP361-203. It describes the data to be summarized and analyzed, including specifics of the statistical analyses to be performed.

This statistical analysis plan (SAP) is based on protocol version 6.00, dated 18-Jun-2019 which incorporates amendment 5.00. Hereafter, this protocol version is referred to as the Clinical Study Protocol (CSP).

# 2. STUDY OBJECTIVES

## 2.1. PRIMARY OBJECTIVE

To evaluate the efficacy of flexibly dosed SEP-363856 (25, 50, or 75 mg/day) in subjects with Parkinson's disease psychosis (PDP) as measured by the Scale for Assessment of Positive Symptoms – Parkinson's Disease (SAPS-PD) at Week 6.

## 2.2. SECONDARY OBJECTIVES

Double-blind Period

To evaluate the efficacy of SEP-363856 (25, 50, or 75 mg/day) in subjects with PDP as measured by:

- Neuropsychiatric Inventory (NPI)
- Clinical global Impression-Severity (CGI-S)
- Mini mental state examination (MMSE) for cognition

## 2.3. SAFETY OBJECTIVES

Double-blind Period:

To evaluate the safety and tolerability of SEP-363856 (25, 50, or 75 mg/day) using

- Adverse event (AE) reports
- Clinical laboratory results
- 12-lead electrocardiograms (ECG)

- Vital signs
- Physical and neurological examinations (PE, NE)
- Body weight, body mass index (BMI)
- Columbia – Suicide Severity Rating Scale (C-SSRS)

Open Label Period:

To evaluate the safety and tolerability of open-label flexibly dosed SEP-363856 (25, 50, or 75 mg/day)

- Adverse event (AE) reports
- Clinical laboratory results
- 12-lead electrocardiograms (ECG)
- Vital signs
- Physical and neurological examinations (PE, NE)
- Body weight and body mass index (BMI)
- Columbia – Suicide Severity Rating Scale (C-SSRS)

## 2.4. OTHER OBJECTIVES

Double-blind Period:

- To evaluate the effects of SEP-363856 sleep quality as measured by the Scales for Outcomes in Parkinson's Disease for nighttime sleep quality (SCOPA-NS) and daytime sleepiness (SCOPA-DS).
- To evaluate the effects of SEP-363856 on motor symptoms associated with Parkinson's Disease as measured by the Unified Parkinson's Disease Rating Scale Parts 2 and 3 (UPDRS II and III).
- To evaluate the effects of SEP-363856 on Rapid Eye Movement (REM) Sleep Behavioral Disorder (RBD) symptoms as measured by the RBD questionnaire (RBDQ).
- Open Label Period: To further evaluate the effectiveness of flexibly-dosed SEP-363856 over 12-weeks of treatment using.:

- Neuropsychiatric Inventory (NPI)
- Clinical Global Impression-Severity (CGI-S)
- Mini Mental State Exam (MMSE)
- SAPS-PD

- To evaluate the effects of SEP-363856 sleep quality as measured by the Scales for Outcomes in Parkinson's Disease for nighttime sleep quality (SCOPA-NS) and daytime sleepiness (SCOPA-DS)
- To evaluate the effects of SEP-363856 on motor symptoms associated with Parkinson's Disease as measured by the Unified Parkinson's Disease Rating Scale Parts 2 and 3 (UPDRS II and III).
- To evaluate the effects of SEP-363856 on Rapid Eye Movement (REM) Sleep Behavioral Disorder (RBD) symptoms as measured by the RBD questionnaire (RBDQ).

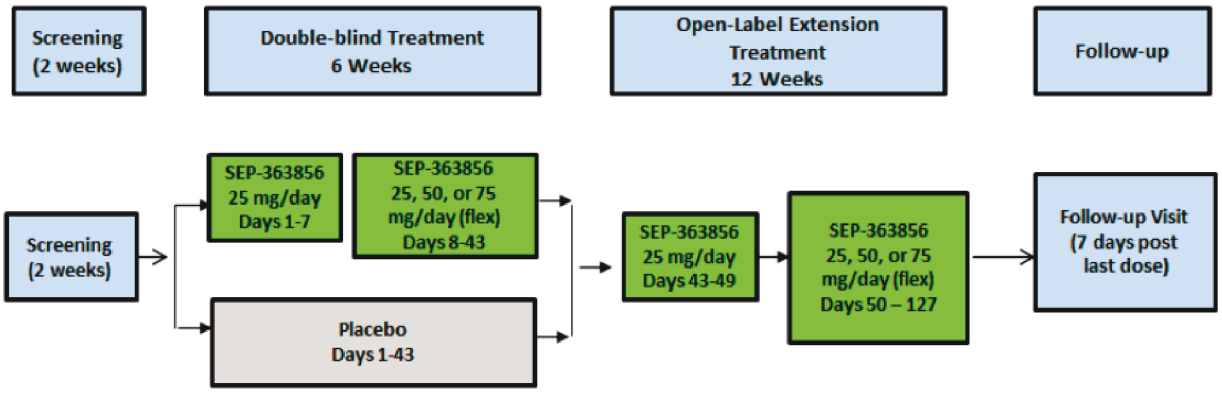
# 3. STUDY DESIGN

## 3.1. GENERAL DESCRIPTION

This is a multicenter, randomized, parallel-group, placebo-controlled study evaluating the efficacy, safety, and tolerability of double-blind SEP-363856 flexibly dosed at 25, 50, or 75 mg/day for 6 weeks followed by 12 weeks of open-label extension of SEP-363856 flexibly-dosed at 25, 50, or 75 mg/day in male and female subjects ≥ 55 years of age with a clinical diagnosis of PDP. The study will randomize approximately 36 subjects to 2 treatment groups in a 2:1 ratio (approximately 24 subjects to SEP-363856 and 12 to placebo)..

The study will consist of 4 periods: Screening/Washout Period (up to 14 days prior to Double-blind Treatment), Double-blind Treatment Period (6 weeks), Open-label SEP-363856 Treatment Period (12 weeks), and Follow-up Period (1 week after last dose) as shown in ***Figure 1: Study Schematic 1***. All post-Baseline clinic visits will have a window of ± 2 days relative to the date of the Baseline visit (Visit 3).

**Figure 1: Study Schematic 1**




## 3.2. METHOD OF ASSIGNING SUBJECTS TO TREATMENT GROUPS

The randomization schedule will be generated by a non-study biostatistician. Once a subject is deemed eligible to be randomized at Day 1 (Visit 3), the Interactive voice/web response system (IXRS) will perform treatment assignment. Subjects will be randomized to one of the following treatment groups for the double-blind period in a 2:1 ratio (SEP-363856:placebo):

- SEP-363856 (25, 50 or 75 mg/day flexible dosing QD for 6 weeks)
- Placebo (QD for 6 weeks)

Once a randomization number has been assigned, it cannot be reused

## 3.3. BLINDING

During the Double-blind Treatment Period, subjects, Investigator staff, persons performing the assessments, clinical operations personnel, data analysts, and personnel at central laboratories will remain blind to the identity of the treatment from the time of randomization until database lock and unblinding, using the following methods; (1) randomization data are kept strictly confidential until the time of unblinding, and will not be accessible by anyone else involved in the study with the exception of bioanalytical personnel involved in the analysis of pharmacokinetic (PK) samples; (2) the identity of the treatments will be concealed by the use of study drugs that are all identical in packaging, labelling, schedule of administration and appearance.

Actual subject identity for plasma concentrations of SEP-363856 and SEP-363854 will not be disclosed before the database lock and the unblinding of the double-blind treatment. After completion of the Double-blind Treatment Period, all subjects will receive unblinded flexibly dosed SEP-363856.

In the case of a medical emergency, where knowledge of study drug by the Investigator or an authorized delegate is essential for immediate medical management, a 24-hour code-break service will be available via the IXRS. The date and reason for unblinding are to be documented in the source documents. Any subject for whom the treatment assignment was unblinded is to be discontinued from further study participation. The subject should return for a final study assessment as described in Section 11.9.9 of the CSP.

## 3.4. DETERMINATION OF SAMPLE SIZE

A total of approximately 36 subjects will be randomized in 2:1 ratio to SEP-363856 and placebo, with approximately 24 subjects assigned to SEP-363856 and 12 subjects assigned to placebo. The sample size will provide powers of 59%, 48%, 37%, and 27% to detect treatment effect sizes of 0.8, 0.7, 0.6, and 0.5, respectively, in change from Baseline in SAPS-PD at Week 6 for SEP-363856 versus placebo. It was estimated by using a two independent sample t-test method with 2-sided significant level of 0.05. The sample size was determined for the purposes of exploring the efficacy, safety and tolerability of flexible dosing with SEP-363856 (25, 50, or 75 mg/day) for 6 weeks in male and female subjects with a clinical diagnosis of PDP. A sufficient number of subjects will be enrolled to ensure 36 randomized subjects.

## 3.5. CHANGES IN THE CONDUCT OF THE STUDY

There is no change in the conduct of the study.

## 3.6. SCHEDULE OF EVENTS

A schedule of events can be found in Section 11 and Table 2: Schedule of Assessments of the CSP.

## 3.7. CHANGES TO ANALYSIS FROM PROTOCOL

There is no change in the analysis from protocol.

# 4. PLANNED ANALYSES

The following analyses will be performed for this study:

- Two database locks and sets of analyses: One after completion of DB phase.
- One after completion of OLE phase.

## 4.1. DATA MONITORING COMMITTEE (DMC)

There will be no planned DMC for this study.

## 4.2. INTERIM ANALYSIS

No interim analysis is planned for this study.

## 4.3. FINAL ANALYSIS

All final, planned analyses identified in this SAP will be performed by IQVIA Biostatistics following Sunovion authorization of this SAP, Sunovion authorization of analysis population, database lock, and unblinding of treatment.

All PK/PD analysis will be described in a separate analysis plan.

# 5. ANALYSIS POPULATIONS

Agreement and authorization of subjects included/ excluded from each analysis set will be conducted prior to the unblinding of the study.

## 5.1. SCREENED [SCR] SUBJECTS

The screened subjects (SCR) set contains of all subjects who provided informed consent for this study.

## 5.2. ENROLLED [ENR] POPULATION

The enrolled (ENR) population contains all screened subjects who passed screening, defined as all of the eligibility criteria met at screening visit.

## 5.3. RANDOMIZED [RND] POPULATION

The all subjects randomized (RND) population contains all subjects in the ENR set who had given their informed consent and were randomized to study medication.

For analyses and displays based on RND, subjects will be classified according to randomized treatment.

## 5.4. MODIFIED INTENT-TO-TREAT [MITT] POPULATION

The modified intention-to-treat (mITT) population is defined as all subjects who were randomized, received at least one dose of study drug, and had a baseline and at least one post-baseline total score in SAPS-PD or NPI or CGI-S during the DB treatment period.

The mITT population will be the primary population for the efficacy analyses. Subjects will be analyzed according to randomized treatment group.

## 5.5. MITT COMPLETE CASES [MITTCOMP1] SAPS-PD POPULATION

The mITT complete cases (mITTComp) is defined as all subjects of the mITT who completed SAPS-PD assessment at the Week 6 visit (Visit 7 not including the ET visit), and have Baseline and Week 6 SAPS-PD total score data available. If a subject had missing SAPS-PD data at visits other than Week 6 (i.e. Visits 3 [Week 1], Visit 4 [Week 2], Visit 5 [Week 3], or Visit 6 [Week 5]), or if he had missing data in assessments other than SAPS-PD at Week 6 (Visit 7), this subject will still be included in the complete case analysis for SAPS-PD total score. Week 6 visit is defined in the protocol as study visit day 43 (+/-2 days). The mITT Complete cases population will be reviewed in a similar manner as the Per-Protocol population as some subjects may have performed a Week 6 visit outside of the visit window but they are still completers.

## 5.6. MITT COMPLETE CASES [MITTCOMP2] CGI-S POPULATION

Same as mITTComp1, for CGI-S instead.

The mITTComp1 and mITTComp2 populations will be used for the complete case analysis of the SAPS-PD total score and CGI-S score respectively. Subjects will be analyzed according to randomized treatment group.

## 5.7. PER PROTOCOL [PP] POPULATION

The per protocol (PP) population comprises all mITT population subjects who:

- Received assigned study medication as randomized
- Have 14 days or more continuous exposure
- Have 75% - 125% compliance, both limit values inclusive
- Have no major protocol violations, determined by a blinded data review prior to database lock

Selected efficacy endpoints will be analyzed using the PP population. Subjects will be analyzed according to randomized treatment group.

## 5.8. SAFETY [SAF] POPULATION

The safety population includes all subjects who were randomized and received at least one dose of study drug during the double-blind period. Safety population will be the primary population for the safety analyses. Subjects will be analyzed according to the actual treatment received (i.e., placebo vs SEP-363856).

## 5.9. OPEN LABEL EXTENSION SAFETY [SAFOL] POPULATION

The OL extension safety population will consist of all subjects who receive at least one dose of study drug during the 12-week open-label extension period. The OL extension safety population will be used for the long-term safety, tolerability, and efficacy analyses. All the by-treatment summaries will be based on the treatment received during the DB period (i.e. SEP-363856to SEP-363856 vs. Placebo to SEP-363856). Changes from baseline in all safety measures will be summarized at each scheduled post-baseline extension visit, based on both the DB baseline and the OL baseline.

# 6. GENERAL CONSIDERATIONS

## 6.1. REFERENCE START DATE AND STUDY DAY

Study Day will be calculated from the reference start dates and will be used to show start/ stop day of assessments and events.

Reference start date is defined as the day of the first dose of study medication during double-blind period, (Day 1 is the day of the first dose of study medication), and will appear in every listing where an assessment date or event date appears.

- If the date of the event is on or after the reference date then:

Study Day = (date of event – reference start date) + 1.

- If the date of the event is prior to the reference date then:

Study Day = (date of event – reference start date).

In the situation where the event date is partial or missing, Study Day, and any corresponding durations will appear missing in the listings. Partial dates will be however presented as is in the listings.

## 6.2. BASELINE

Unless otherwise specified, baseline of DB period is defined as the last non-missing measurement taken prior to reference start date (including unscheduled assessments). Unless otherwise specified, baseline of OL period is defined as the measurement at Visit 7.

Whenever available, the time information should be accounted for in the derivation of baseline values. In the case where time isn't available and the date of last non-missing measurement and the reference start date coincide, that measurement will be considered baseline.

Adverse Events (AEs) and medications commencing on the reference start date will be considered post-baseline.

## 6.3. DERIVED TIME POINTS

For subjects with missing SAPS-PD total score at Week 6 (Visit 7), a last observation carried forward (LOCF) endpoint for SAPS-PD total score will be derived (for use in sensitivity analysis), using the last post-baseline value up to the Week 6 visit date. Similarly, missing CGI-S score, missing NPI score, missing NPI (H+D) score and missing MMSE score are also imputed using the same LOCF method.

Both scheduled and unscheduled assessments that are collected post baseline will contribute to the derivation of LOCF endpoint.

LOCF imputation will be performed on data with the early termination data already mapped (see Section 6.4).

## 6.4. RETESTS, UNSCHEDULED VISITS AND EARLY TERMINATION DATA

In general, for by-visit summaries, data recorded at the scheduled visits will be presented.

Unscheduled measurements will not be included in by-visit summaries. Unscheduled measurements collected prior to the first dose of study medication will contribute to the derivation of the baseline value. Unscheduled measurements collected post baseline will contribute to the derivation of the LOCF endpoint value, markedly abnormal post-baseline value, and best/ worst case value where required

(e.g. shift tables).

In the case of a retest (same visit number assigned), the latest available measurement for that visit will be used for by-visit summaries.

Early termination data collected post baseline will be assigned to the next planned visit for that assessment. This mapping will be done for all data points used in the efficacy and safety analyses, except for the complete case analysis of the SAPS-PD total score and CGI-S score (see Section 15.1.4.5).

Listings will include scheduled, unscheduled, retest and early termination data.

## 6.5. WINDOWING CONVENTIONS

All data will be analyzed according to the schedule outlined in the CSP and according to the visit denoted on the case report form (CRF). Apart from early termination data (see Section 6.4), no visit windowing will be performed during the analysis for this study.

## 6.6. STATISTICAL TESTS

All statistical inference, unless otherwise stated, will be performed with 2-sided tests at the significance level of 0.05 and 2-sided 95% confidence intervals (CIs), unless otherwise specified in the description of the analyses.

## 6.7. TREATMENT GROUPS

There are several doses administrated in this flexibly-doses study but statistical analysis will describe/analyze Placebo group versus SEP-363856 all doses pooled. Apart from some exposure tables, no analysis will be performed by doses received.

## 6.8. STUDY MEDICATION

In this SAP, study medication refers to double-blind medication as well as open-label medication.

## 6.9. SUBJECTS ON-PERIOD

The SAPS-PD, CGI-S, MMSE, UPDRS II, UPDRS III, and neurological exam should be performed during the subject's on-periods (i.e., at a time when the subject shows no clinical evidence of worsening in their Parkinson's symptoms, as determined by the site investigator). Also, the NPI should be completed referencing the subject's on-period.

Regarding those scales, for statistical analysis, only the subjects on-period assessment will be taken

into account. Any assessment not done during subjects on-period will be considered as missing for statistical analysis. For data listings, all values will be listed and the variable subjects on-period will also be presented.

### 6.10. COMMON CALCULATIONS

For quantitative measurements, change from baseline will be calculated as:

- Value at visit X – baseline value

Percentage change from baseline will be calculated as:

- (Value at visit X – baseline value)*100 / (baseline value)

### 6.11. SOFTWARE VERSION

All analyses will be conducted using SAS version 9.4 or higher.

## 7. STATISTICAL CONSIDERATIONS

### 7.1. ADJUSTMENTS FOR COVARIATES AND FACTORS TO BE INCLUDED IN ANALYSES

The following covariates are used in the analyses. For details of their inclusion in the models, see specific analysis sections.

- Baseline value of the variable to be analyzed

### 7.2. MULTICENTER STUDIES

This study will be conducted by multiple investigators at multiple sites in the US. No adjustment for geographic region is needed.

Center pooling will not be carried out for use in analyses for this study.

### 7.3. MISSING DATA

Missing safety data will not be imputed.

Missing efficacy data will be handled as described in Section 16.1.2 of this SAP.

For rating scales, such as SAPS-PD and CGI-S, unless otherwise specified, if any item score contributing to the total/subscale score is missing, then the total /subscale score will be set to missing.

## 7.4. MULTIPLE COMPARISONS/ MULTIPLICITY

There is only one primary efficacy comparison of change from baseline in SAPS-PD total score at Week 6. Therefore, no adjustment for multiplicity is planned in this study. For all other analyses, nominal p-values will be presented; no multiplicity adjustment will be performed.

## 7.5. EXAMINATION OF SUBGROUPS

Subgroup analyses will be conducted as stated in the respective efficacy and safety analysis sections. It should be noted that the study was not powered to detect treatment differences within subgroups.

The following subgroups will be assessed:

- Sex:
  - Female
  - Male
- Age group (years):
  - >=55 to <65
  - >=65 to <75
  - >=75
- Number of prior hospitalizations for PDP:
  - 0
  - 1
  - 2
  - 3 or more
- Duration of PDP (years):
  - < 5
  - >= 5 to < 10
  - >= 10 to < 20
  - >= 20
- Baseline MMSE (categories for subgroup analysis are different than subgroups presented for the demographic summaries, see Section 10):
  - >24
  - <=24

- Baseline NPI Hallucinations + Delusions (NPI [H+D]) Severity:
  - >= 6 to < 12
  - >=12

- Lead-in Status:
  - Lead-in
  - Non Lead-in

# 8. OUTPUT PRESENTATIONS

APPENDIX 1 shows conventions for presentation of data in outputs.

The templates provided with this SAP describe the presentations for this study and therefore the format and content of the summary tables, figures and listings to be provided by IQVIA Biostatistics.

# 9. DISPOSITION AND WITHDRAWALS

All subjects who provide informed consent will be accounted for in this study.

For DB period, subject disposition will be presented for the SCR set and described by the randomized DB treatment group (where applicable) and overall for all subjects. The number and percentage of subjects who were screened, screen-failed, enrolled, randomized, received double-blind study medication, and completed or discontinued early from the DB treatment phase (including reasons for discontinuation), will be presented. The number of subjects who discontinued from the study before or at a given visit will be described by the treatment group and overall on the RND subjects.

For OL period, subject disposition will be presented for the RND set. The number and percentage of subjects who entered OL period, received at least one dose of OL study medication, and completed or discontinued early from the OL treatment phase (including reasons for discontinuation), will be presented. The number of subjects who discontinued from the study before or at a given visit will be described on the RND subjects.

Important protocol deviations (IPDs) will be identified and documented based on blinded reviews of data listings for DB period and on unblinded reviews of data listing for OL period. The IPD categories may include, but may not be limited to:

- Did not satisfy important inclusion, exclusion, and/or randomization criteria (DB period and Overall period)
- Received any disallowed concomitant medication
- Overall double-blind compliance rate < 75% or > 125% (DB period and OL period)

Further details on the identification of IPDs are provided in the Important Protocol Deviation Review Specifications document.

Individual IPDs will be presented for all randomized subjects in a data listing.

For the DB and OLE periods, IPDs will be identified for all randomized subjects and presented in data listings. The number and percentage of subjects within each IPD category will be summarized by treatment group and overall for the mITT populations.

# 10. DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

For DB period, demographic data and other baseline characteristics will be presented for the mITT population, safety population and PP population. For mITT and PP population, the data will be presented by randomized treatment groups and overall. For safety population, the data will be presented by the actual treatment received and overall. Basic demographic data will also be summarized for SCR population and for subjects enrolled by randomization status (i.e. randomized versus not randomized).

For OL period, demographic data and other baseline characteristics will be presented for the SAFOL population.

No statistical testing will be carried out for demographic or other baseline characteristics.

Categorical demography variables will be summarized by presenting the number and percentage subjects in each category. Continuous demography variables will be summarized using descriptive statistics (number of subjects, mean and standard deviation [SD], median, 25th percentile [Q1], 75th percentile [Q3], minimum, and maximum). The following demographic and other baseline characteristics will be reported for this study:

- Age (years) - calculated relative to date of informed consent, as a continuous variable and categorically:
    - >=55 to <65
    - >=65 to <75
    - >=75
- Sex
    - Female
    - Male
- Race
    - American Indian or Alaska Native
    - Asian
    - Black or African American
    - Native Hawaiian or Other Pacific Islander
    - White
    - Multiracial
    - Other

- Ethnicity
    - Hispanic or Latino
    - Not Hispanic or Latino
- Weight (kg) as a continuous variable
- Height (cm) as a continuous variable
- BMI ($kg/m^2$), as a continuous variable and categorically:
    - Underweight: <18.5
    - Normal: 18.5 to <25.0
    - Overweight: 25.0 to <30.0
    - Obese: >=30.0
- Baseline SAPS-PD total score, SAPS-PD Hallucinations subscale score, SAPS-PD Delusions subscale score as continuous variables and categorically:
    - < Overall median value at baseline
    - >= Overall median value at baseline
      Note: Median value for the above categories is based on subjects from the SAF population.
- Baseline NPI total score and NPI subscale scores (Hallucinations, Delusions, Agitation, Depression, Anxiety, Hallucinations + Delusions [H+D]) as continuous variables and NPI (H+D) score categorically:
    - >= 6 to < 12
    - >=12
- Baseline MMSE Score:
    - <=16
    - 17-19
    - 20-24
    - 25-27
    - 28-30
- Baseline CGI-S score, as a continuous variable and categorically:
    - < 4
    - >= 4

Weight, BMI, Baseline SAPS-PD and Baseline CGI-S for OL is defined as the last nonmissing value

prior to the first dose of study drug in OL period.

The following psychiatric history data will be summarized only for DB period for the mITT population, safety population, and PP population in a separate table.

- Time since initial onset of PDP, in years, calculated relative to date of informed consent; both as a continuous variable and categorically (See APPENDIX 2 for partial date imputation rules for initial onset of PDP)
    - < 5
    - >= 5 to < 10
    - >= 10 to < 20
    - >= 20
- Number (0, 1, 2, 3 or more) of prior hospitalizations for treatment of PDP.

## 10.1. DERIVATIONS

- Time since initial onset of Parkinson disease (years):

  (Date of informed consent – date of initial onset of PD+1) / 365.25

- Time since initial onset of PDP (years):

  (Date of informed consent – date of initial onset of PDP+1) / 365.25

- BMI ($kg/m^2$) = weight (kg)/height $(m)^2$

# 11. MEDICAL AND SURGICAL HISTORY

Medical and surgical history information, including both past and concomitant medical conditions and major surgical history, as collected on the Medical History CRF form, will be coded using Medical Dictionary for Regulatory Activities (MedDRA), Version 19.0 or higher, and presented for DB period by treatment group for the safety population by System Organ Class (SOC) and Preferred Term (PT).

# 12. MEDICATIONS

Medications will be presented for double-blind period and open label-period based on the double-blind safety population and open-label safety population respectively and coded using the WHO drug dictionary, Version 01JUN2016E or higher.

Whenever available, the time information should be accounted for in the derivation of prior, concomitant, and post-treatment medications. See APPENDIX 2 for the handling of partial dates for medications. In the case where it is not possible to define a medication as prior, concomitant, or post-treatment, the

medication will be classified by the worst case; i.e. concomitant.

For DB period,

- Prior medications are medications which stopped prior to the first dose of DB study medication.
- Summary of antipsychotic use prior to screening (total and %), by type: nuplazid, quietapine (Seroquel), clozapine, other
- Concomitant medications are medications which started at the same time of or after the first dose of DB study medication and at the same time of or before the last dose of DB study medication; or started prior to and ended at the same time of or after the first dose of DB study medication; or started at the same time of or prior to the last dose of DB study medication and marked as ongoing.
- Post-treatment medications are medications which started after the last dose of DB study medication.

For OL period,

- Prior medications are medications which stopped prior to the first dose of OL study medication.
- Concomitant medications are medications which started at the same time of or after the first dose of OL study medication and at the same time of or before the last dose of OL study medication; or started prior to and ended at the same time of or after the first dose of OL study medication; or started at the same time of or prior to the last dose of OL study medication and marked as ongoing.
- Post-treatment medications are medications which started after the last dose of OL study medication.

Prior and concomitant medication use will be summarized by Anatomical Therapeutic Chemical (ATC) Level 3 classification and preferred name using frequencies and percentages. Prior, concomitant, and post-treatment medications will be provided in data listings.

# 13. STUDY MEDICATION EXPOSURE

Duration of exposure to study medication will be summarized for the safety population.

Interruptions, compliance, and dose changes are not taken into account for duration of exposure.

Duration (in days) of exposure will be summarized both as a continuous variable for double-blind treatment period and open-label extension period, and categorically:

For double-blind:

- Number and percentage of subjects with drug exposure ≥ 1, ≥ 4, ≥ 7, ≥ 14, ≥ 21, ≥ 28, ≥ 35, and ≥ 42 days;

- Number and percentage of subjects with drug exposure for 1 - 3, 4 - 6, 7 - 13, 14 – 20, 21 – 27, 28 – 34, 35 – 41 and ≥ 42 days.

For open-label:

- Number and percentage of subjects with open-label extension study drug exposure ≥ 1, ≥ 7, ≥ 14, ≥ 21, ≥ 28, ≥ 56, and ≥ 84 days;

- Number and percentage of subjects with extension study drug exposure for 1 - 6, 7 - 13, 14 – 20, 21 – 27, 28 – 55, 56 – 83 and ≥ 84 days.

The modal daily dose during the double-blind period and the open-label period (i.e. the dose level that a subject was on for the most number of days during the study) will be summarized categorically by treatment group:

- 25 mg/day
- 50 mg/day
- 75 mg/day
- Tie between 25 mg/day and 50 mg/day
- Tie between 25mg/day and 75 mg/day
- Tie between 50mg/day and 75 mg/day
- Tie between 25mg/day, 50 mg/day, and 75 mg/day

Include histogram showing distribution of each dose by visit. Include linear plots for each subject showing dose taken over time.

For DB period, the number of days that a subject was on the 25 mg/day dose level, on the 50 mg/day dose level and on the 75 mg/day dose level will also be summarized, both as a continuous variable and categorically:

- 1 – 3 days
- 4 – 6 days
- 7 – 13 days
- 14 – 20 days
- 21 – 27 days
- 28 – 34 days
- 35 – 41 days
- ≥ 42 days

For OL period, the number of days that a subject was on the 25 mg/day dose level, on the 50 mg/day dose level and on the 75 mg/day dose level will also be summarized, both as a continuous variable and categorically:

- 1 – 6 days
- 7 – 13 days
- 14 – 20 days

- o 21 – 27 days
- o 28 – 55 days
- o 56 – 83 days
- o ≥ 84 days

Mean daily dose will be calculated for the DB period and entire open-label extension period as the cumulative dose (mg) of SEP-363856 divided by the duration of exposure (in days), where cumulative dose is the sum of all doses a subject received during the DB period or open-label extension period. Modal daily dose will be determined as the daily dose that is taken for the most time (in terms of number of days) among all doses taken during the DB or OL extension period. Both mean daily dose and modal daily dose will be summarized.

The dose adjustment decision at each visit will be summarized in a shift table for DB and for OL extension period, for the mITT and safety populations.

## 13.1. DERIVATIONS

- Duration of exposure (in days) will be calculated as:

last dose date – first dose date + 1

For double-blind period, first dose date and last dose will be taken from the eCRF “Study Drug Administration / Drug Accountability - Double Blind” form.

For open-label period, first dose date will be taken from the eCRF “Study Drug Administration / Drug Accountability – Open-label” form. Last dose date will mainly be taken from the EX/DA page, which should match the eCRF “End of Study” form.

# 14. STUDY MEDICATION COMPLIANCE

Compliance to study medication will be summarized for the safety population by treatment group and overall for DB period. Compliance to study medication will be summarized for the open-label safety population for OL extension period.

Percent compliance will be calculated by visit and overall for the double-blind and open-label period.

Non-compliance is defined as less than 75% or more than 125% non-missing compliance for the double-blind period and OL extension period. Subjects with missing compliance will not be classified as non-compliant.

Compliance will be summarized both as a continuous variable (i.e. mean percentage) and categorically (i.e. number and percentage of subjects who are compliant vs. non-compliant, or with compliance < 75%, 75% - 125%, > 125%, and missing).

## 14.1. DERIVATIONS

Compliance with study medication will both be calculated for each period defined by study visits (i.e. per-visit compliance) and overall.

Per-visit compliance for the period defined by visit (V-1) and visit V will be calculated as:

$$\frac{\#\ Capsules\ dispensed\ at\ Visit\ (V\text{-}1)\ -\ \#\ Capsules\ returned\ at\ Visit\ V\ -\ \#\ Capsules\ reported\ lost\ during\ this\ period}{\#\ Capsules\ should\ be\ taken\ per\ day \times (Date\ of\ Visit\ V\ -\ Date\ of\ Visit\ (V\text{-}1))} \times 100\%$$

One capsule per day are supposed to be taken. If any of the following numbers are missing, per-visit compliance will not be calculated for the period impacted: number capsules dispensed at Visit (V-1), number capsules returned at Visit V, and number capsules reported lost during this period.

If a subject discontinued from the study in between Visit (V-1) and Visit V, then the above formula will be modified to:

$$\frac{\#\ Capsules\ dispensed\ at\ Visit\ (V\text{-}1)\ -\ \#\ Capsules\ returned\ at\ Visit\ ET\ -\ \#\ Capsules\ reported\ lost\ during\ this\ period}{\#\ Capsules\ should\ be\ taken\ per\ day \times (Date\ of\ Visit\ ET\ -\ Date\ of\ Visit\ (V\text{-}1))} \times 100\%$$

The per-visit compliance calculated above will contribute to the summary of per-visit compliance for the period defined by visit (V-1) and visit V.

Overall compliance will be calculated as:

$$\frac{Total\ \#\ capsules\ dispensed\ -\ Total\ \#\ capsules\ returned\ -\ Total\ \#\ capsules\ reported\ lost}{\#\ Capsules\ should\ be\ taken\ per\ day \times Duration\ of\ Exposure} \times 100\%$$

Duration of exposure is calculated as specified in Section 14.1.

If any of the following numbers are missing at one or more visits, overall compliance will not be calculated: number capsules dispensed, number capsules returned, number capsules reported lost.

# 15. EFFICACY OUTCOMES

In addition to the analyses described in Sections 15.1.3, 15.1.4, 15.2.3, 15.3.3, all efficacy variables will be summarized descriptively by treatment group and visit (including the Week 6 LOCF endpoint) for the mITT population. For SAPS-PD total score and CGI-S score, the data will also be summarized descriptively for each subgroup.

## 15.1. PRIMARY EFFICACY

### 15.1.1. PRIMARY EFFICACY VARIABLE & DERIVATION

The primary efficacy variable is the change from baseline in SAPS-PD total score at Week 6 in DB period for testing superiority of SEP-363856 to placebo.

SAPS-PD is a Parkinson's disease-adapted subset of nine items derived from the Scale for Assessment of Positive Symptoms (SAPS), including seven items assessing individual symptoms (four items for hallucinations and three items for delusions), a global hallucinations item and a global delusions item. Separate items are rated from 0 (absent) to 5 (severe), for a total possible score on the SAPS-PD ranging from 0 to 45. Total score will be equal to the sum of the seven items plus the global hallucinations, plus the global delusions.

The SAPS-PD subscale for hallucinations is defined as the sum of the 4 items for hallucinations and the global hallucination item. The SAPS-PD subscale for delusions is defined as the sum of the three items for delusions and the global delusion item.

There can be three different SAPS-PD item results: "Original" SAPS-PD item result, "Review" SAPS-PD item result, and "Consensus" SAPS-PD item result. When calculating the SAPS-PD total score or subscale scores, the following rule will be used when selecting individual items to calculate the total score or subscale scores:

1) Select the "Consensus" item result when available.
2) If the "Consensus" item result is not available, then select the "Original" item result. The "Review" item score should never be used to calculate the SAPS-PD total score or subscale scores.

The choice of "Consensus" or "Original" item result is selected independently for each of the seven items from the SAPS-PD questionnaire.

For the above total score and subscale scores, if any item contributing to its calculation is missing then the score will be set to missing. Only subjects on-period data will be taken into account in this analysis (see Section 6.9)

### 15.1.2. Missing Data Methods for Primary Efficacy Variable(s)

The SAPS-PD total score will be set to missing if any one item is missing. The primary efficacy variable, change from baseline in SAPS-PD total score at Week 6, will be set to missing if SAPS-PD total score at Week 6 is missing. The same applies to all other visits.

The primary analysis of the primary efficacy variable will use a mixed model for repeated measures (MMRM) based on observed data (with early termination data mapped as described in Section 6.4). Missing data will not be imputed.

The analysis of covariance (ANCOVA) of the primary efficacy variable will utilize the Week 6 LOCF endpoint as sensitivity analysis, which is imputed as described in Section 6.3.

### 15.1.3. Primary Analysis of Primary Efficacy Variable(s)

The primary objective of this study is to test whether the mean changes from baseline in SAPS-PD total score at Week 6 in the SEP-363856 (25, 50 or 75 mg/day) group and the placebo group are equal.

Let $\mu_{SEP}$ and $\mu_{PBO}$ represent the mean changes from baseline at Week 6 in SAPS-PD total score for the SEP-363856 (25, 50 or 75 mg/day) and placebo groups, respectively. The following hypothesis will be tested:

**$H_0$:** $\mu_{SEP} = \mu_{PBO}$ versus $H_1$: $\mu_{SEP} \neq \mu_{PBO}$

The primary efficacy analysis will be performed on the mITT population.

Change from baseline in SAPS-PD total score at Week 6 will be analyzed using an MMRM model, with fixed factors for visit (Weeks 2, 3, 5 and 6; as a categorical variable), treatment, and treatment-by-visit interaction, and with baseline SAPS-PD total score as a covariate. An unstructured covariance matrix will be used to model the within-subject correlation. The Kenward-Roger approximation will be used to calculate the denominator degrees of freedom.

In case the model above fails to converge, a spatial exponential covariance structure and a spatial power covariance structure will be assumed sequentially. The first covariance structure to yield convergence will be used in the analysis.

Within group effect size at each time point will be calculated as the least squares (LS) mean of each treatment group divided by the model estimate of standard deviation, obtained as the square root of the corresponding diagonal element of the residual covariance matrix (R matrix from PROC MIXED). Between group effect size at each time point will be calculated as the LS mean difference divided by the model estimate of standard deviation, obtained as described above.

LS means (±SE) of change from baseline in SAPS-PD total score will be plotted by week and treatment group.

The above analysis will also be performed on the SAPS-PD subscale score for hallucinations and the SAPS-PD subscale score for delusions.

### 15.1.4. SENSITIVITY ANALYSIS OF PRIMARY EFFICACY VARIABLE(S)

#### 15.1.4.1. Graphical examination

Patterns of missing data in the primary endpoint will be assessed through graphical tools based on the reason and timing of study withdrawal by subjects.

Change from baseline in SAPS-PD total score by visit will be described and plotted by reasons of early discontinuation (adverse event, lack of efficacy, progressive disease, withdrawal by subject, protocol deviation, noncompliance with study medication, death, pregnancy, other reasons) and for completers, separately for each treatment group. Similar reasons may be combined, depending on the number of subjects under each reason.

Change from baseline in SAPS-PD total score by visit will also be described and plotted by time of termination (Week 1 terminators, Week 2 terminators, Week 3 terminators, Week 5 terminators, Week 6 terminators) and for completers, separately for each treatment group.

- Week 1 terminators: subjects who discontinued before or on Visit 3;
- Week 2 terminators: subjects who discontinued after Visit 3 but before or on Visit 4;
- Week 3 terminators: subjects who discontinued after Visit 4 but before or on Visit 5;
- Week 5 terminators: subjects who discontinued after Visit 5 but before or on Visit 6;
- Week 6 terminators: subjects who discontinued after Visit 6 but before or on Visit 7;
- Completers: subjects who discontinued after Visit 7.

#### 15.1.4.2. Pattern-mixture model with placebo-based multiple imputation

The MMRM model used in the primary analysis makes the assumption that data are missing at random (MAR). However, the missing data mechanism may or may not be at random. Sensitivity to the missing data assumptions will be tested by using the pattern-mixture model with placebo-based multiple imputation method (Ratitch, O'Kelly, & Tosiello, 2013), exploring the robustness of the MMRM results of the primary efficacy analysis. In this analysis, missing values in the SEP-363856 (25, 50 or 75 mg/day) treatment group will be imputed based on data of the placebo group, assuming that, after withdrawal, subjects from the SEP-363856 (25, 50 or 75 mg/day) group will exhibit the same future evolution of Parkinson's as subjects from the placebo group, and that subjects who discontinue from the placebo group will exhibit the same future evolution of Parkinson's as subjects in the placebo group remaining in the study. This approach does not assume a sustained benefit of experimental treatment after discontinuation.

Two separate imputation procedures are used to impute missing values. Firstly, the Markov chain Monte Carlo (MCMC) method is used to perform partial imputation to obtain datasets with monotone missing patterns. Then a sequential regression multiple imputation method is used to impute the monotone missing values.

Under the assumption that the SAPS-PD total scores have a multivariate normal distribution, the MCMC method is used to impute only intermittent missing values (using the SAS MI procedure with MCMC statement), by using a data augmentation algorithm, with each iteration n consisting of an imputation step and a posterior step. The imputation step uses a random draw of $\theta^{(n)}$, parameter of the joint imputation model, to sample missing values from a conditional distribution $P(Y_{mis}|x, y_{obs}, \theta^{(n)})$, obtaining $y_{mis}^{(n)}$, the subset of missing values that need to be filled in to achieve monotone missingness. The posterior step simulates a new draw of the parameter $\theta^{(n+1)}$ from the posterior distribution given the current monotone missing data $P(\theta|x, y_{obs}, y_{mis}^{(n)})$ with a non-informative Jeffreys prior. Treatment group will be taken into account for the imputation. These steps are repeated to obtain 1000 datasets with monotone missingness. The options for burn-in and thinning will be optimized by examining the diagnostic plots. The random seed number is 12345.

The remaining monotone missing data will be imputed using sequential regression multiple imputation, where a separate regression model is estimated for imputation of each variable (i.e., SAPS-PD total score at each time point). Imputation of values in the placebo group will assume MAR. Imputation of values in the SEP-363856 (25, 50 or 75 mg/day) group will be done as if the subject had been a member of the placebo group. Missing values in the SEP-363856 (25, 50 or 75 mg/day) group will be imputed using the imputation model of the placebo group, i.e., conditional on subject values observed at time points prior to discontinuation relative to the mean of the model for the placebo group. Each sequential regression model (i.e., for imputation of values at a given time point) will include explanatory variables for all previous (Baseline, Week 1, 2, 3 and 5) values of SAPS-PD total score. Missing values at a given time point in placebo and SEP-363856 (25, 50 or 75 mg/day) arms will be imputed from the same imputation model, conditional on subject values observed or imputed at previous time points. The SAS MI procedure with the MONOTONE REG statement is used to specify that the regression method will be used for the imputation, and the MNAR statement with MODEL option will be used for the SAPS-PD total score at each post-baseline visit to specify that only observations from the placebo group should be used to estimate the imputation model. The random seed number is 12345.

No rounding restriction will be applied to imputed SAPS-PD total scores. The imputed SAPS-PD total scores must be within the range of 0 to 45.

Each of the 1000 imputed datasets will be analyzed using the same MMRM model as the primary efficacy analysis. Results from the analysis of each imputed dataset, i.e. the LS means of each treatment group, the LS mean treatment difference, and their standard errors, will be combined using Rubin's imputation rules (using the SAS MIANALYZE procedure) to produce pooled LS mean estimates, their standard errors and 95% CI, and a pooled p-value for the test of null hypothesis of no treatment effect.

#### 15.1.4.3. Tipping point analysis

Sensitivity to departures from the MAR assumption will also be investigated using a tipping point analysis. In this analysis, departures from MAR in the SEP-363856 (25, 50 or 75 mg/day) group will be assessed assuming that subjects who discontinue the study have, on average, efficacy outcomes after discontinuation that are worse by some amount δ compared to other similar subjects with observed data at the same time point (i.e., compared to a value which would have been assumed under a MAR model).

A series of analyses will be performed with increasing values of δ until the analysis conclusion of a statistically significant treatment effect no longer holds. The value of δ that overturns the primary results will represent a tipping point. An interpretation of clinical plausibility of the assumption underlying the tipping point will be provided.

Change from baseline in SAPS-PD total score will be analyzed based on data observed while the subject remains on study as well as data imputed using multiple imputation methodology for the time points at which no value is observed. Intermittent (non-monotone) missing data will be imputed first based on the MAR assumption and a multivariate joint Gaussian imputation model using the MCMC method within each treatment arm, as described above for the pattern-mixture model with placebo-based multiple imputation.

The remaining monotone missing data will be imputed using sequential regression multiple imputation, where a separate regression model is estimated for imputation of each variable (i.e., SAPS-PD total score at each time point). Each regression model will include explanatory variables for treatment and all previous (Baseline, Week 2, 3, and 5) values of SAPS-PD total score. After the MAR-based imputations have been generated for SAPS-PD total score at each time point, a value of δ will be added to all imputed values in the SEP-363856 (25, 50 or 75 mg/day) group. This approach assumes that the marginal mean of unobserved subject measurements is worse by δ at each time point after discontinuation compared to the marginal mean of subjects with observed data at the same time point.

No rounding restriction will be applied to imputed continuous values. The imputed SAPS-PD total scores must be within the range of 0 to 45.

One-thousand (1000) imputed datasets will be generated. The random seed number for both the partial imputation step and the sequential regression multiple imputation step will be 12345.

Each of the 1000 imputed and δ-adjusted datasets will be analyzed using the same MMRM model as the primary efficacy analysis. Results from the analysis of each imputed dataset, i.e. the LS mean treatment difference and its standard error, will be combined using Rubin's imputation rules (using the SAS MIANALYZE procedure) to produce a pooled LS mean estimate of treatment difference, its standard error and 95% CI, and a pooled p-value for the test of null hypothesis of no treatment effect.

Analyses will be conducted with different values of δ at each visit, which represents a percentage of the LS mean treatment difference at that visit, starting at 5% with 5% increments, until either the tipping point is identified or the 100% penalty is applied.

#### 15.1.4.4. Analysis of covariance

As another supportive analysis for the efficacy endpoints of change from Baseline in SAPS-PD total score at Week 6, missing data at Week 6 will be imputed using the last observation carried forward (LOCF) method, and the data will be analyzed using an ANCOVA model. The response variable for the model will be: change from Baseline in SAPS-PD total score at the Week 6 LOCF endpoint. The ANCOVA model will include treatment as fixed factors and include Baseline score as a covariate.

The proportion of subjects by percent change from baseline at the LOCF Endpoint will be plotted by treatment group.

#### 15.1.4.5. Complete case analysis

The complete case analysis will be performed on the mITT complete case population (see Section 5.5).

The MMRM analysis used in the primary efficacy analysis will be repeated on the above subjects.

#### 15.1.4.6. Analysis on the per protocol population

Sensitivity to the analysis population will be tested by repeating the MMRM analysis used in the primary efficacy analysis on the PP population.

### 15.1.5. SUBGROUP ANALYSIS OF PRIMARY EFFICACY VARIABLE(S)

For each of the subgroup factors listed in Section 7.5, change from baseline in SAPS-PD total score will be analyzed using the MMRM method. The MMRM model will include fixed effects for treatment, subgroup, visit, baseline SAPS-PD total score, and treatment-by-subgroup, treatment-by-visit, subgroup-by-visit, and treatment-by-subgroup-by-visit interactions.

The estimates obtained from the MMRM models will be presented separately for each subgroup. The p-value for the treatment-by-subgroup interaction at Week 6 will be presented. Its statistical significance will be assessed at the 0.10 level for homogeneity of the treatment effect across the different categories of a subgroup factor. In case a significant interaction effect is detected, estimates by subgroup will be examined to determine the nature of the interaction (qualitative or quantitative).

### 15.1.6. ANALYSIS OF PRIMARY EFFICACY VARIABLE FOR OPEN-LABEL PERIOD(S)

The observed values of the primary efficacy measures at the OL baseline, and at each scheduled post-baseline open-label extension visit, will be summarized descriptively. Changes from open-label baseline in these measures will be summarized at each scheduled post-baseline open-label extension visit, based on both the DB and the OL baseline.. .

## 15.2. SECONDARY EFFICACY

The analyses of the secondary efficacy variables will be performed on the mITT population.

### 15.2.1. SECONDARY EFFICACY VARIABLES & DERIVATIONS

15.2.1.1. Change from baseline in CGI-S score at Week 6

The CGI-S is a clinician-rated assessment of the subject's current illness state on a 7-point scale, where a higher score is associated with a greater illness severity. The CGI-S score takes one of the following values: 1 (normal, not at all ill), 2 (borderline mentally ill), 3 (mildly ill), 4 (moderately ill), 5 (markedly ill), 6 (severely ill), 7 (among the most extremely ill patients). Only subjects on-period data will be taken into account in this analysis (see Section 6.9).

15.2.1.2. Change from baseline in NPI score and NPI (H+D) score at Week 6

The NPI (Cummings-1994) is a 12-item behavior rating scale composed of a structured interview of the caregiver, which assess psychiatric disturbance. The 12 sub-domains are the following:

- Delusions
- Hallucinations
- Agitation\Aggression
- Depression\Dysphoria
- Anxiety
- Elation\Euphoria
- Apathy\Indifference
- Disinhibition
- Irritability\Lability
- Aberrant Motor behavior
- Sleep and Nighttime behavior disorders
- Appetite and eating changes

For each sub-domain, a gating question is asked:

- If the responses to these questions indicate that the patient has problems with a particular subdomain of behavior, the caregiver is only then asked all the questions about that domain, rating the frequency of the symptoms on a 4-point scale and their severity on a 3-point scale (Cummings-1997). The sub-domain score is the product of the frequency score multiplied by the severity score for that behavioral domain, it will be calculated by statistical programming, the results of the eCRF will not be used. In case of any missing data for frequency or severity then the sub-domain score will be set to missing.

- If the responses to these questions indicate that the patient does not have problems with that particular subdomain of behavior, that sub-domain score will be set to zero.

For each sub-domain, a measure of the level of caregiver distress is also given but is not included in the NPI score and not used in any statistical analysis.

A NPI score is obtained by summing all the sub-domain scores. Only subjects on-period data will be taken into account in this analysis (see Section 6.9).

For the combined score NPI (H+D) = NPI hallucinations + NPI delusions, if either domain is missing, the combined score will be set to the non-missing domain score. If both domain scores are missing, then

the combined score will be set to missing.

#### 15.2.1.3. Change from baseline in MMSE score at Week 6

The MMSE for Cognition (Folstein-1975) is a brief instrument, used to assess cognitive function, consisting of 11 tests including orientation, memory (recent and immediate), concentration, language, and praxis. Scores range from 0 to 30, with lower scores indicating greater cognitive impairment. The total score will be calculated by statistical programming, the results of the eCRF will not be used. Only subjects on-period data will be taken into account in this analysis (see Section 6.9).

#### 15.2.1.4. Proportion of subjects who achieve a SAPS-PD response

SAPS-PD response is defined as a 30%, 50% and 100% or greater improvement (i.e. decrease) in SAPS-PD total score from Baseline.

The percent change in SAPS-PD total score from Baseline will be calculated by:

$$\frac{\text{SAPS-PD total score at a visit or the LOCF endpoint - SAPS-PD total score at Baseline}}{\text{SAPS-PD total score at Baseline}} \times 100\%$$

For each subject, the responder indicator will be set to 1 if the percent change is <= -30%. The indicator will be set to 0 if the percentage is > -30%. The indicator will be set to missing if the percentage is missing.

In addition, SAPS-PD response defined by a series of additional thresholds will be assessed; that is, having a 50% or greater and 100% or greater improvement in SAPS-PD total score from Baseline.

SAPS-PD response at all three thresholds will be derived for all post-baseline time points during the double-blind period.

#### 15.2.1.5. Subjects with SAPS-PD percent change from baseline

The proportion of subjects achieving a given SAPS-PD percent change from baseline or lower at the Week 6 LOCF endpoint will be calculated for each treatment group. This calculation will be performed at multiple levels of percent change from baseline, from -100% to >=100%, with 5% increments. The results will be reported in a graph with the percent change from baseline threshold on the x-axis and proportion of subjects in each treatment group on the y-axis.

### 15.2.2. Missing Data Methods for Secondary Efficacy Variable(s)

For rating scales with more than one item, if any item score contributing to the total score or individual domain score is missing, then the total score or individual domain score will be set to missing. The change from baseline scores will also be set to missing accordingly.

The MMRM analysis of the continuous secondary efficacy variables will be based on observed data (with early termination data mapped as described in Section 6.4). Missing data will not be imputed. SAPS-PD responder calculation will be based on the SAPS-PD total score at the Week 6 LOCF endpoint.

The ANCOVA analyses of the continuous secondary efficacy variables will utilize the change from baseline at the Week 6 LOCF endpoint, as a sensitivity analysis. Refer to section 6.3 for further details on LOCF imputation.

### 15.2.3. Analysis of Secondary Efficacy Variables

#### 15.2.3.1. Analysis of change from baseline in CGI-S score at Week 6 during the DB period

Change from baseline in CGI-S score at Week 6 will be analyzed using an MMRM model similar to the model used in the primary analysis of the primary efficacy variable (see Section 15.1.3), with baseline CGI-S score as a covariate. Only subjects on-period data will be taken into account in this analysis (see Section 6.9).

Relationships between CGI change and SAPS-PD will be explored.

#### 15.2.3.2. Analysis of change from baseline in NPI score and NPI (H+D) score at Week 6 during the DB period

Change from baseline in NPI score at Week 6 will be analyzed using an MMRM model similar to the model used in the primary analysis of the primary efficacy variable (see Section 15.1.3), with baseline NPI score as a covariate. Only subjects on-period data will be taken into account in this analysis (see Section 6.9). These summaries on NPI will be repeated on NPI (H+D).

Relationships between NPI and NPI (H+D) changes and SAPS-PD will be explored.

#### 15.2.3.3. Analysis of change from baseline in MMSE score at Week 6 during the DB period

Change from baseline in MMSE score at Week 6 will be analyzed using an MMRM model similar to the model used in the primary analysis of the primary efficacy variable (see Section 16.1.3), with baseline MMSE score as a covariate (>24 or <= 24). Only subjects on-period data will be taken into account in this analysis (see Section 6.9)

### 15.2.4. Sensitivity Analysis of Secondary Efficacy Variables

Sensitivity analyses of CGI-S score will be performed similarly to those outlined in Section 15.1.4.

Sensitivity analyses of NPI and MMSE score will be performed using the ANCOVA model described in Section 15.1.4.4.

### 15.2.5. Subgroup Analysis of Secondary Efficacy Variable(s)

Subgroup analysis of CGI-S score will be performed similarly to those described in Section 15.1.5.

### 15.2.6. Analysis of SASPS-PD Responder

SAPS-PD response will be analyzed using a logistic regression model with responder indicator as the dependent variable and treatment as categorical factors and baseline SAPS-PD total score as a covariate. The analysis will be performed at the Week 6 LOCF endpoint as well as for each of the post-baseline scheduled visits during the double-blind period.

Odds ratio and its 95% CI as estimated by the model will be presented. In addition, the Number Needed to Treat (NNT) will be calculated for SEP-363856 group as:

$$\text{NNT} = \frac{1}{\text{Risk Reduction (RR)}} = \frac{1}{\text{SAPS-PD Response Rate}_{\text{SEP-363856}} - \text{SAPS-PD Response Rate}_{\text{Placebo}}}$$

The 95% CI of NNT will be obtained by taking the reciprocal of the 95% CI bounds of the Absolute Risk Reduction when both lower and upper confidence limits are positive. The NNT and its 95% CI will be provided as integers; any fractional values will be rounded up to the nearest integer (the lower confidence limit will be rounded down and the upper confidence limit will be rounded up)

### 15.2.7. Analysis of Secondary Efficacy Variable for Open-label period(s)

The observed values of the secondary efficacy measures at the OL baseline, and at each scheduled post-baseline open-label extension visit, will be summarized descriptively. Changes from open-label baseline in these measures will be summarized at each scheduled post-baseline open-label extension visit, based on both the DB and the OL baseline.

## 15.3. Other Efficacy Variables

### 15.3.1. Other Efficacy Variables & Derivations

15.3.1.1. Parkinson's Disease Sleep Scale (SCOPA) Nighttime Sleep (NS) / Daytime Sleepiness (DS)

SCOPA is a self-administered, validated short questionnaire that is used to assess NS problems (SCOPA-NS) and DS (SCOPA-DS) in subjects with Parkinson's disease.

The NS subscale addresses NS problems in the past month and includes 5 items with 4 response options ranging from 0 (not at all) to 3 (a lot). The maximum score of this subscale is 15, with higher scores reflecting more severe sleep problems. One additional question evaluates overall sleep quality

on a 7-point scale (ranging from slept very well to slept very badly). The score on this item is not included in the score of the NS scale but is used separately as a global measure of sleep quality.

The DS subscale evaluates DS in the past month and includes 6 items with 4 response options, ranging from 0 (never) to 3 (often). The maximum score is 18, with higher scores reflecting more severe sleepiness (Marinus-2003).

#### 15.3.1.2. Unified Parkinson's disease Rating scale (UPDRS) parts II and III

The UPDRS is a scale used to follow the longitudinal course of Parkinson's disease. From the original scale, only parts II and III will be performed in this study.

For UPDRS Part II (Activities of Daily Living), 13 items will be recorded from 0 to 4. It will represent historical information from the past seven days as how the patient functioned in various activities of daily living. The maximum score is 52.

For the UPDRS Part III (Motor Examination), 25 items will be recorded from 0 to 4. It will represent an objective motor assessment at the time of evaluation. The maximum score is 100.

All items from both parts are scored on a 5-point scale (0 - 4) with higher scores indicating higher severity. Only subjects on-period data will be taken into account in this analysis (see Section 6.9).

#### 15.3.1.3. Rapid Eye Movement (REM) Sleep Behavior Disorder Questionnaire (RBDQ)

The RBDQ is a 13-item patient self-rated questionnaire addressing the clinical manifestations of REM Sleep Behavior Disorder (RBD). Questions are brief and answered "yes" or "no", with a maximum total score of 13.

Items 1 - 4 assess the frequency and content of dreams and their relationship to nocturnal movements and other behavior. Item 5 inquires about self-injuries and injuries to the bed partner. Item 6 (including 4 subitems) assesses specific nocturnal motor behaviors, items 7 and 8 address nocturnal awakenings, item 9 assesses disturbed sleep in general and item 10 queries the presence of any neurological disorder (Stiasny-Kolster-2007, Cummings-2014).

#### 15.3.1.4. Zarit Burden Interview (Zarit-22)

The Zarit Burden Interview was designed to reflect subjective burden of caregivers of impaired elderly patients. The caregiver is asked to complete a questionnaire of 22 questions about the impact of the patient's disability on their life. For the first 21 items the caregiver is to indicate on a 5-point scale (0 - 4) how often they felt that way (Never, Rarely, Sometimes, Quite Frequently, or Nearly Always). On the 22nd item the caregiver indicates their overall impression of burden (Not at All, A Little, Moderately, Quite a Bit, Extremely). Zarit-22 questionnaire was administered during versions 1 though 5 of the protocol and was removed from version 6 of the protocol.

### 15.3.2. MISSING DATA METHODS FOR OTHER EFFICACY VARIABLE(S)

For rating scales with more than one item, if any item score contributing to the total score or individual domain score is missing, then the total score or individual domain score will be set to missing. The change from baseline scores will also be set to missing accordingly.

The ANCOVA analyses of the continuous other efficacy variables will utilize the change from baseline at the Week 6 LOCF endpoints.

### 15.3.3. ANALYSIS OF OTHER EFFICACY VARIABLES

15.3.3.1. Analysis of change from baseline in SCOPA-NS and SCOPA-DS total score at Week 6

The SCOPA-NS and SCOPA-DS total score and change from baseline will be summarized by timepoint for each treatment group. For SCOPA-NS and SCOPA-DS, change from baseline at Week 6 will be analyzed by an ANCOVA with treatment as fixed effect and baseline SCOPA-NS (or SCOPA-DS) as covariate. The ANCOVA analysis will utilize the Week 6 LOCF endpoint, which is imputed the same way the imputation done for primary endpoint (see Section 6.3).

15.3.3.2. The Overall sleep quality will be described (categorical) at each timepoint by treatment group.Analysis of change from baseline in UPDRS II and III total score at Week 6

The UPDRS II and III total score and change from baseline will be summarized by timepoint for each treatment group. Change from baseline at Week 6 will be analyzed by an ANCOVA with treatment as fixed effect and baseline UPDRS II or III total score as covariate. The ANCOVA analysis will utilize the Week 6 LOCF endpoint, which is imputed the same way the imputation done for primary endpoint (see Section 6.3). Only subjects on-period data will be taken into account in this analysis (see Section 6.9).

15.3.3.3. Analysis of change from baseline in RBDQ total score at Week 6

The RBDQ total score and change from baseline will be summarized by timepoint for each treatment group. Change from baseline at Week 6 will be analyzed by an ANCOVA with treatment as fixed effect and baseline RBDQ total score as covariate. The ANCOVA analysis will utilize the Week 6 LOCF endpoint, which is imputed the same way the imputation done for primary endpoint (see Section 6.3).

15.3.3.4. Displays of Zarit Burden Interview (Zarit-22)

Items from the Zarit-22 questionnaire collected during previous versions of the protocol will be presented in a data listing.

## 16. SAFETY OUTCOMES

All outputs for safety outcomes will be based on the safety population.

There will be no statistical comparisons between the treatment groups for safety data, unless otherwise specified in the relevant sections. Descriptive statistics will be presented for anxiolytic or hypnotic medication taken since last visit.

## 16.1. ADVERSE EVENTS

Both adverse events (AEs) and pre-treatment events will be coded using MedDRA central coding dictionary, Version 18.1 or higher.

For all adverse events analysis, first dose of study medication is defined as the first dose during DB period and OL period respectively.

Adverse events are untoward medical occurrences that started at the same time of or after the first dose of study medication. Untoward medical occurrences that started between ICF and prior to the first dose of study medication are pre-treatment events, i.e. events collected at Visit 1, Visit 2 and Visit 3.

Whenever available, the time information should be accounted for in the derivation of adverse events vs. pre-treatment events. In the case where time isn't available, untoward medical occurrences that started on or after the day of the first dose of study medication will be considered adverse events; those that started before the day of the first dose of study medication will be considered pre-treatment.

For the purpose of statistical analysis, all AEs in the DB period which started on or after the first dose of DB period study medication and within 9 days after the last dose of DB period study medication (given that this is before the first dose of the OL period study medication for subjects entering into the OL period), or had a partial or missing start date such that it cannot be determined whether the adverse event is within the 9-day window, will be included in DB period AE table summaries. All AEs in the OL period which started on or after the first dose of OL period study medication and within 9 days after the last dose of OL period study medication, or had a partial or missing start date such that it cannot be determined whether the adverse event is within the 9-day window, will be included in OL period AE table summaries.

See APPENDIX 2 for handling of partial dates for adverse events. In the case where it is not possible to define an untoward medical occurrence as adverse event or pre-treatment event, it will be classified by the worst case, i.e. adverse event.

An overall summary of the incidence of adverse events for DB period, OL extension period within each of the categories described in the following sections will be provided as specified in the templates. This summary for DB period will also be repeated by the race, sex, age, number of prior hospitalizations for PDP, and duration of PDP subgroups. The overall incidence summary for DB period, OL extension period and overall period will also be provided for AEs related to study medication.

Listings for DB period, OL extension period will be provided for all AEs, AEs leading to discontinuation of study medication, AEs leading to discontinuation from the study, serious adverse events (SAE), and AEs leading to death. A listing for pre-treatment events will also be presented.

For incidence summaries, each subject will be counted only once within each SOC and PT. If not otherwise specified, all summaries will present incidence (number of subjects and percentages) and number of events.

### 16.1.1. ALL AES

AEs will be presented by SOC, High Level Term (HLT), and PT for AE incidence and number of events for DB period, OL extension period. These will be presented by frequency and dose at onset. Listings

for each subject will be included. A listing of all AEs will be presented for DB period, OL extension period.

AEs will also be presented by maximum severity and by strongest relationship to the study medication as specified in the sections below for DB period, OL extension period.

AEs that occurred in >= 5% of subjects in either treatment group will be summarized by SOC and PT for DB period, OL extension period.

#### 16.1.1.1. Severity

Severity is classed as mild/ moderate/ severe (increasing severity). AEs with a missing severity will be summarized as missing severity. If a subject reports an AE more than once within the same SOC/ PT, the AE with the worst severity will be used in the corresponding severity summaries. For this summary, AEs will be presented by SOC and PT for DB period, OL extension period.

#### 16.1.1.2. Relationship to Study Medication

Relationship, as indicated by the investigator, is classed as “not related”/ “possible”/ “probable”/ “definite” (increasing strength of relationship). A dichotomous approach will be utilized for the safety output: a “related” AE is defined as an AE with a relationship to the study medication of “possible”, “probable” or “definite”. A “not related” AE is defined as an AE with a relationship to the study medication of “not related”. AEs with a missing relationship to the study medication will be regarded as “related” to the study medication. If a subject reports the same AE more than once within the same SOC/ PT, the AE with the strongest relationship to study medication will be used in the corresponding relationship summaries. For this summary, AEs will be presented by SOC and PT for DB period, OL extension period.

### 16.1.2. AEs Leading to Discontinuation of Study Medication

AEs leading to permanent discontinuation of study medication are AEs for which “Drug Withdrawn” is selected for “Action Taken with Study Treatment” on the AE CRF page. A summary of AEs leading to discontinuation of study medication by SOC and PT will be presented for DB period and OL extension period, including dose at time of treatment discontinuation. A listing of AEs leading to discontinuation of study medication will be presented for DB period and OL extension period.

### 16.1.3. AEs Leading to Discontinuation from the Study

AEs leading to discontinuation from the study are AEs with “Caused Study Discontinuation” = “Yes” on the AE CRF page. A summary of AEs leading to discontinuation from the study by SOC and PT will be presented for DB period and OL extension period, including dose at time of early discontinuation. A listing of AEs leading to discontinuation from the study will be presented for DB period and OL extension period.

### 16.1.4. SERIOUS ADVERSE EVENTS

SAEs are those AEs recorded as "Serious" on the AE CRF page. Summaries of serious AEs and non-serious AEs by SOC and PT will be prepared for DB period and OL extension period. A listing of SAEs will be presented for DB period and OL extension period.

### 16.1.5. ADVERSE EVENTS LEADING TO DEATH

AEs leading to death are those AEs which are recorded as having an outcome of "Fatal" on the AE CRF page. A summary of AEs leading to death by SOC and PT will be prepared for DB period and OL extension period. A listing of deaths will be presented for DB period and OL extension period.

### 16.1.6. ADVERSE EVENTS BY SUBGROUP

As stated above, overall incidence summaries will be presented by the subgroups of race, sex, age, prior number of hospitalizations for PDP, the duration of PDP, and MMSE. The same subgroup factors will also apply to the by-subgroup summaries for the following events:

- All AEs, by SOC and PT

In addition, all AEs (by SOC and PT) will be summarized by the baseline BMI ($kg/m^2$) category:

- Underweight: <18.5
- Normal: 18.5 to <25.0
- Overweight: 25.0 to <30.0
- Obese: >=30.0

## 16.2. LABORATORY EVALUATIONS

Laboratory data to be reported for this study include hematology, serum chemistry (including lipid panel and thyroid panel), urinalysis, coagulation, urine drug screening, Glycosylated haemoglobin (HbA1c), serum follicle stimulating hormone (FSH), prolactin, breath alcohol test and serum pregnancy test (only listed). A list of laboratory assessments to be included in the outputs is included in Section 21/ Appendix II of the CSP.

Presentations will use international system of units (SI).

Quantitative laboratory measurements reported as "< X", i.e. below the lower limit of quantification (BLQ), or "> X", i.e. above the upper limit of quantification (ULQ), will be converted to X for the purpose of quantitative summaries, but will be presented as recorded, i.e. as "< X" or "> X" in listings.

The following summaries will be provided for laboratory data for DB and OL period:

- By visit summary of observed values and changes from baseline for continuous data in hematology, chemistry, and urinalysis. Prolactin results will be summarized separately by gender. Glucose and lipid panel results will be summarized separately by fasting status.

- By visit summary of the number and percentage of subjects in each outcome category for categorical data in urinalysis and urine drug screening (if applicable). For urine drug screening, the results will be reported as “Positive”/ “Negative”.

- Shift in lab results (chemistry, hematology, urinalysis) from baseline to Week 6 in DB period, and from baseline to post-OL baseline visits in OL period according to the reference range criteria provided by the central laboratory.

For the OL extension period, both the DB baseline and the OL baseline will be used in the calculation of change from baseline values, and for shift from baseline.

All laboratory data will be provided in data listings for DB period and OL period, with the values outside the reference ranges and that meet potentially clinically significant (PCS) criteria shown in Appendix 4 flagged.

## 16.3. ECG EVALUATIONS

Results from the central ECG (Electrocardiogram) will be included in the reporting of this study.

The following ECG parameters will be reported for this study:

- PR Interval (msec)
- RR Interval (msec)
- QRS Interval (msec)
- QT Interval (msec)
- QTcF Interval (msec) [derived]
- QTcB Interval (msec) [derived]
- Heart rate (HR) (beats/min)
- ECG findings
- Overall assessment of ECG (investigator’s judgment):
    - Normal
    - Abnormal, clinically Significant [CS]
    - Abnormal, not clinically significant [NCS]

The following summaries will be provided for ECG data for DB period and OL period:

- By visit summary of observed values and changes from baseline (for quantitative measurements)

- By visit summary of ECG overall assessment results
- Shift in ECG overall assessments from baseline to Week 6 for DB period, and from baseline to post-OL baseline visits for OL period

Number and percentage of subjects with QTc levels in each of the QTc categories will be presented. For the OL extension period, both the DB baseline and the OL baseline will be used in the calculation of change from baseline values, and for shift from baseline.

The number and percentage of subjects with QTc values in the following categories will be identified for DB period and OL period. The same criteria apply to both QTcF and QTcB.

- > 450 msec at any post-baseline time point (including unscheduled visits) not present at baseline
- > 480 msec at any post-baseline time point (including unscheduled visits) not present at baseline
- > 500 msec at any post-baseline time point (including unscheduled visits) not present at baseline
- ≥ 30 msec and <60 msec increase from baseline for at least one post-baseline measurement (including unscheduled visits)
- ≥ 60 msec increase from baseline for at least one post-baseline measurement (including unscheduled visits)

All ECG parameters, overall interpretation, and findings will be provided in data listings.

### 16.3.1. ECG Specific Derivations

The following three measures are provided by the ECG vendor and those provided values will be analyzed. If one of these values is missing and derivation in the course of analysis is needed, the following derivations will be used:

- Bazett's Correction of QT interval (msec)
  - $QTcB\ (\text{msec}) = \frac{QT\ (ms)}{\sqrt{RR\ (ms)/1000}}$
- Fridericia's Correction of QT interval (msec)
  - $QTcF\ (\text{msec}) = \frac{QT\ (ms)}{\sqrt[3]{RR\ (ms)/1000}}$

- RR Interval – If RR Interval is not available it will be derived from HR as follows, for the derivation of the QTc corrections
  - $RR\ (\text{msec}) = 1000 * \frac{60}{HR\ (bpm)}$

## 16.4. VITAL SIGNS

The following vital signs measurements will be reported for this study:

- Supine Systolic Blood Pressure (mmHg)
- Standing Systolic Blood Pressure (mmHg)
- Supine Diastolic Blood Pressure (mmHg)
- Standing Diastolic Blood Pressure (mmHg)
- Supine Pulse Rate (beats/min)
- Standing Pulse Rate (beats/min)
- Respiratory Rate (breaths/min)
- Temperature (C)
- Weight (kg)
- BMI ($kg/m^2$)

The following summaries will be provided for vital signs data:

- Observed value and change from baseline by visit including the follow-up visit for DB period and OL period. For the OL extension period, both the DB baseline and the OL baseline will be used in the calculation of change from baseline values
- BMI Categories will be described by visit for DB period and OL period.

All vital signs data will be provided in a data listing for DB period and OL period, including flagging of PCS results.

**PCS Criteria**

| Parameter Name | PCS Low | PCS High |
|---|---|---|

| | | |
|---|---|---|
| Systolic Blood Pressure (mmHg) Standing and Supine | Value ≤ 90 and ≥ 20 decrease from baseline | Value ≥ 180 and ≥ 20 increase from baseline |
| Diastolic Blood Pressure (mmHg) Standing and Supine | Value ≤ 50 and ≥ 15 decrease from baseline | Value ≥ 105 and ≥ 15 increase from baseline |
| Pulse Rate (beats/min) Standing and Supine | Value ≤ 50 and ≥ 15 decrease from baseline | Value ≥ 120 and ≥ 15 increase from baseline |
| Weight (kg) | ≥ 7% decrease from baseline | ≥ 7% increase from baseline |
| Temperature (°C) | Value <= 35°C | Value ≥ 38.3°C |
| Respiratory (breaths/min) | Value ≤ 11 and ≥ 10 decrease from baseline | Value ≥ 25 and ≥ 10 increase from baseline |

### 16.4.1. ORTHOSTATIC EFFECTS

Orthostatic hypotension is defined as a decrease of >= 20 mmHg in systolic blood pressure or >= 10 mmHg in diastolic blood pressure after the subject had been standing for at least 2 to 4 minutes, compared to the systolic blood pressure and diastolic pressure measured in the supine position, respectively.

Orthostatic tachycardia is defined as a heart rate increase of >= 20 beats per minute (bpm) and a heart rate of >100 bpm after the subject was standing for at least 2 to 4 minutes, compared to the heart rate measured in the supine position.

The number and percentage of subjects with orthostatic hypotension and orthostatic tachycardia will be summarized for baseline and the overall post-baseline period including follow-up visit, as well as by visit for DB period and OL period. This will be presented for systolic hypotension, for diastolic hypotension, and for both, including mean change from baseline, any orthostatic hypotension or tachycardia events that occurred at the early termination visit will be assigned to the next planned visit (as specified in Section 6.4).

## 16.5. PHYSICAL AND NEUROLOGICAL EXAMINATION

Physical and neurological examination will be provided in a data listing including follow-up data. Subjects on-period information will be added to the neurological examination listing.

## 16.6. OTHER SAFETY ASSESSMENTS

### 16.6.1. COLUMBIA SUICIDE SEVERITY RATING SCALE (C-SSRS)

The C-SSRS is a tool designed to systematically assess and track suicidal behavior and suicidal ideation for life time, one month (30 days) prior to the screening visit, and throughout the study. The strength of this suicide classification system is in its ability to comprehensively identify suicidal events while limiting the over-identification of suicidal behavior. The C-SSRS Baseline/Screening Version is used at the screening visit and the C-SSRS Since Last Visit Version is used from Visit 2 (lead-in) onward. Subjects with Type 4 (active suicidal ideation with some intent to act, without specific plan) or Type 5 (active suicidal ideation with specific plan and intent) suicidal ideation during the study will be discontinued from the study and referred to a mental health professional.

C-SSRS includes four sections: Suicidal Ideation, Intensity of Ideation, Suicidal Behavior, and Answer for Actual Suicide Attempts.

The following outcomes are C-SSRS categories and have binary responses (yes/no). The categories are re-ordered from the scale to facilitate the definitions of the C-SSRS endpoints, and to provide clarity in the presentation of the results.

Suicidal ideation is measured by 5 categories, representing 5 subtypes of suicidal ideation with increasing severity:

- Category 1: Wish to be Dead
- Category 2: Non-specific Active Suicidal Thoughts
- Category 3: Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act
- Category 4: Active Suicidal Ideation with Some Intent to Act, without Specific Plan
- Category 5: Active Suicidal Ideation with Specific Plan and Intent

Suicidal behavior is measured by 5 categories, representing 5 subtypes of suicidal behavior:

- Category 6: Preparatory Acts or Behavior
- Category 7: Aborted Attempt
- Category 8: Interrupted Attempt
- Category 9: Actual Attempt (non-fatal)
- Category 10: Completed Suicide

The 10 categories above are not mutually exclusive. Subjects will be counted in each category for which they have an event.

Self-injurious behavior without suicidal intent is a non-suicide-related C-SSRS outcome, and also has a binary response (yes/no).

For the purpose of C-SSRS analysis, “baseline” and “post-baseline” for DB period are defined as follows.

| Time point | Study Visit | C-SSRS Version | Derivation Rule |
|---|---|---|---|
| Baseline | Screening/Visit 1 | Baseline/Screening – Past 1 Month | Most severe outcome |

| | | | |
|---|---|---|---|
| | All visits occurring prior to the first dose of study medication during double-blind period | Since Last Visit | |
| Post-baseline | All post-baseline visits up to and including Week 6/Visit 7, including unscheduled visits | Since Last Visit | Most severe outcome |

For OL period, “baseline” and “post-baseline” are defined as follows.

| Time point | Study Visit | C-SSRS Version | Derivation Rule |
|---|---|---|---|
| OL Baseline | Week 4/Visit 7 of study 361-201 | Since Last Visit | Most severe outcome |
| OL Post-baseline | All OL post-baseline visits from Week7/Visit 8 up to and including Week 18/Visit 13 and follow-up visit, including unscheduled visits | Since Last Visit | Most severe outcome |

C-SSRS composite endpoints will be derived for each time point of interest (i.e. baseline, post baseline, and each study visit) for DB period and OL period as follows:

- Any suicidal ideation: A “yes” answer to any one of the 5 suicidal ideation questions on C-SSRS (Categories 1-5).
- Any suicidal behavior: A “yes” answer to any one of the 5 suicidal behavior questions on the C-SSRS (Categories 6-10).
- Any suicidality: A “yes” answer to any one of the 10 suicidal ideation and behavior questions on the C-SSRS (Categories 1-10).

For each subject, the suicidal ideation score at each time point of interest (i.e. baseline, post baseline, each study visit including follow-up visit, and the Week 6 LOCF endpoint) is defined as the maximum suicidal ideation category (1-5) present for the time of interest. If no ideation is present a score of 0 is assigned. A suicidal ideation score of 4 or 5 is considered serious.

The number and percentage of subjects with any suicidality, any suicidal ideation and subtypes of ideation, any suicidal behavior and subtypes of behavior, and any non-suicidal self-injurious behavior will be presented for DB period and OL period for:

- Baseline (as defined above)
- Post-baseline (as defined above)

Each scheduled study visits for DB period related: Screening (lifetime; past 1 month), Day 1/Visit 3, Week 2/Visit 4, Week 3/Visit 5, Week 5/Visit 6, Week 6/Visit 7, Week 6 LOCF.

Each scheduled study visits for OL period related: Week 7/Visit 8, Week 8/Visit 9, Week 9/Visit 10, Week 10/Visit 11, Week 14/Visit 12, Week 18/Visit 13 and Follow-up/Visit 14.

Difference between treatment groups for any post-baseline suicidality, suicidal behavior, and suicidal ideation in DB period will be evaluated using Fisher's Exact test.

Shift in suicidal ideation score from baseline to the post-baseline time point and to each of the following study visits will be presented by treatment: Week 2/Visit 4, Week 3/Visit 5, Week 5/Visit 6, Week 6/Visit 7, Week 6 LOCF for DB period. Shift in suicidal ideation score from OL baseline to the OL post-baseline time point and to each of the following study visits will be presented by treatment: Week 8/Visit 9, Week 9/Visit 10, Week 10/Visit 11, Week 14/Visit 12, Week 18/Visit 13 and Follow-up/Visit 14 for OL period

Intensity of ideation for the most severe ideation subtype is measured in terms of frequency, duration, controllability, deterrents, and reasons for ideation. Each is measured with responses ranging from 0 to 5 for frequency and duration, and from 1 to 5 for controllability, deterrents, and reasons for ideation. The ideation intensity total score is the sum of responses to the five items and can range from 2 to 25 for subjects with endorsed suicidal ideation. If one or more of these five items are missing at an assessment, the total score will be set to missing. If a subject did not endorse any suicidal ideation, a score of 0 for the ideation intensity total score will be given.

Actual lethality associated with actual attempts is rated on a 6-point scale from 0 = 'No physical damage or very minor physical damage' to 5 = 'Death'. Potential lethality of actual attempts (if actual lethality = 0) is rated on a 3-point scale from 0 = 'Behavior not likely to result in injury' to 2 = 'Behavior likely to result in death despite available medical care.

The ideation intensity total score and the actual lethality and potential lethality of actual attempts will be presented in data listings for DB period and OL period.

# 17. PHARMACOKINETIC AND PHARMACOGENOMIC ANALYSIS

## 17.1. POPULATION PHARMACOKINETIC ANALYSIS

All plasma concentrations of SEP-363856 and SEP-363854 will be presented in data listings.

Population pharmacokinetic (PK) analysis will be performed using plasma SEP-363856 concentrations. The results will be reported separately.

## 17.2. PHARMACODYNAMIC ANALYSIS

The relationship between SAPS-PD total score and plasma SEP-363856 exposure using population PK/pharmacodynamics (PD) methods will be explored. The results will be reported separately.

### 17.3. PHARMACOGENOMIC ANALYSIS

The impact of cytochrome (CYP) P450 CYP2D6 metabolizer status on plasma SEP-363856 exposure will be explored. The results will be reported separately.

## 18. DATA NOT SUMMARIZED OR PRESENTED

The data not summarized or presented are:

- Any data other than disposition, demographics and adverse events that are collected on screen failures and on enrolled subjects at the previously failed screenings.

These data will not be summarized or presented, but will be available in the clinical study database, SDTM and/or ADaM datasets.

## 19. CHANGES IN THE ANALYSIS SPECIFIED IN THE STATISTICAL ANALYSIS PLAN

Any changes or deviations during the analysis and reporting process from the statistical analysis plan designed will be described and justified in the final report.

## 20. REFERENCES

Ratitch, B. O'Kelly, &Tosiello. (2013). Missing data in clinical trials: from clinical assumptions to statistical analysis using pattern mixture models. . *Pharmaceutical Statistics*.

# APPENDIX 1. PROGRAMMING CONVENTIONS FOR OUTPUTS

## OUTPUT CONVENTIONS

Where applicable, the Appendix_Compilation_Working_Guidelines_Final 07May2014 .pdf document – provided by Sunovion – will be followed.

In addition, the following output conventions are to be followed:

- The first row in the body of the table or listing should be blank
- The left-hand column should start in column 1. No indenting or centering of the output should occur.
- Rounding should be done with the SAS function ROUND.
- Numbers in tables should be rounded, not truncated.
- Alphanumeric output should be left aligned.
- Numbers should be decimal point aligned.
- Whole numbers should be right aligned.
- Text values should be left aligned.
- The first letter of a text entry should be capitalized.
- The width of the entire output should match the linesize (134)

- Univariate Statistics:

  - If the raw data has N decimal places, then the summary statistics should have the following decimal places:
    - Minimum and maximum: N
    - Mean, median, Q1, and Q3: N + 1
    - SD: N + 2

- Frequencies and percentages (n and %):

  - Percent values should be reported inside parentheses, with one space between the count (n) and the left parenthesis of the percentage. Parentheses should be justified to accept a maximum of 100.0 as a value and padded with blank space if the percent is less than 100.0.
  - Percentages will be reported to one decimal place, except cases where percent <100.0% but >99.9% will be presented as '>99.9%' (e.g., 99.99% is presented as >99.9%); and cases where percent < 0.1% will be presented as '<0.1%' (e.g., 0.08% is presented as <0.1%). Rounding will be applied after the <0.1% and >99.9% rule.
  - Where counts are zero, no percentage should appear in the output.

- Confidence Intervals:

  - Confidence intervals and estimates are presented to one place more than the raw data, and standard errors to two places more than the raw data.
  - Confidence intervals should be justified so that parentheses displayed on consecutive lines of a table "line up".
  - Boundary values of confidence intervals should be separated by a comma.

- Boundary values should be padded as necessary to accept negative values and to allow alignment of the decimal place.

- P-values:
  - P-values should be reported to three decimal places, except values <1.000 but >0.999 will be presented as '>0.999' (e.g., 0.9998 is presented as >0.999); and values <0.001 will be presented as '<0.001' (e.g., 0.0009 is presented as <0.001). Rounding will be applied after the <0.001 and >0.999 rule.

- Ratios:
  - Ratios should be reported to one more decimal place than the raw data.

- Spacing:
  - There must be a minimum of 1 blank space between columns (preferably 2).

- Missing values:
  - A "0" should be used to indicate a zero frequency.
  - A blank will be used to indicate missing data in an end-of-text table or subject listing.

- Figures:
  - Figures should be provided in RTF files using the SAS Output Delivery System (ODS), as Computer Graphics Metafile (CGM) formatted graphical output generated by SAS.
  - The CGM file itself should contain the title or footer.
  - The image should be clear and of high quality when viewed in the Word document, and when printed.
  - In general, boxes around the figures should be used.

- Footers should be defined as follows:
  - A continuous line of underscores ('_') will follow the body of the table or listing prior to any footnotes at the bottom of the page.
  - Table footnotes should be defined using compute statements in the proc report, and should appear directly after the body of the table.
  - If text wraps across more than one line (for a note), the first letter for all lines of text after the first one will be indented to align beneath the first letter of the text in the first line.

## DATES & TIMES

Depending on data available, dates and times will take the form yyyy-mm-ddThh:mm:ss.

## SPELLING FORMAT

English US.

## PRESENTATION OF TREATMENT GROUPS

For outputs, treatment groups will be represented as follows and in that order:

| Treatment Group | For Tables, Graphs and Listings in DB period | For Tables, Graphs and Listings in OL period |
|---|---|---|
| Placebo | Placebo | |
| SEP-363856 (25, 50 or 75 mg/day) | SEP-363856 | SEP-363856 |

## LISTINGS

All listings will be ordered by the following (unless otherwise indicated in the template):

- Actual treatment received, displaying SEP-363856 first and then placebo,
- Subject ID,
- Date/Time (where applicable) - listings of adverse events, concomitant medications, medical histories etc. should be sorted in chronological order, with earliest adverse event, medication or history coming first,
- For listings where non-randomized subjects are included, these will appear in a category after the randomized treatment groups labelled ‘Not Randomized’.

# APPENDIX 2. PARTIAL DATE CONVENTIONS

Imputed dates will not be presented in the listings. In the algorithms for date, 361-203 study med start date the date of first dose of study medication during double blind period.

## ALGORITHM FOR ADVERSE EVENTS:

The concept of "date" below should also include time information whenever available.

| START DATE | STOP DATE | ACTION |
|---|---|---|
| Known | Known | If start date < 361-203 study med start date, then pre-treatment events<br>If start date >= 361-203 study med start date, then 361-203 AE |
| Known | Partial | If start date < 361-203 study med start date, then pre-treatment events<br>If start date >= 361-203 study med start date, then 361-203 AE |
| Known | Missing | If start date < 361-203 study med start date, then pre-treatment events<br>If start date >= 361-203 study med start date, then 361-203 AE |
| | | |
| Partial, but known components show that it cannot be on or after 361-203 study med start date | Known | Pre-treatment events |
| Partial, but known components show that it cannot be on or after 361-203 study med start date | Partial | Pre-treatment events |
| Partial, but known components show that it cannot be on or after 361-203 study med start date | Missing | Pre-treatment events |
| | | |

| START DATE | STOP DATE | ACTION |
|---|---|---|
| Partial | Partial | Impute stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown).<br>Then:<br>If stop date < 361-203 study med start date, then pre-treatment events<br>If stop date >= 361-203 study med start date, then 361-203 AE |
| | | |
| Partial, could be on or after 361-203 study med start date | Known | If stop date < 361-203 study med start date, then pre-treatment events<br>If stop date >= 361-203 study med start date, then 361-203 AE |
| Partial, could be on or after 361-203 study med start date | Partial | Impute stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown).<br>Then:<br>If stop date < 361-203 study med start date, then pre-treatment events<br>If stop date >= 361-203 study med start date, then 361-203 AE |
| Partial | Missing | Assumed 361-203 AE |
| | | |
| Missing | Known | If stop date < 361-203 study med start date, then pre-treatment events<br>If stop date >= 361-203 study med start date, then 361-203 AE |

| START DATE | STOP DATE | ACTION |
|---|---|---|
| Missing | Partial | Impute stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown).<br>Then:<br>If stop date < 361-203 study med start date, then pre-treatment events<br>If stop date >= 361-203 study med start date, then 361-203 AE |
| Missing | Missing | Assumed 361-203 AE |

## ALGORITHM FOR PRIOR / CONCOMITANT / POST-TREATMENT MEDICATIONS:

The concept of "date" below should also include time information whenever available.

| START DATE | STOP DATE | ACTION |
|---|---|---|
| Known | Known | If stop date < 361-203 study med start date, assign as prior.<br>If stop date >= 361-203 study med start date and start date <= end of treatment, assign as concomitant.<br>If stop date >= 361-203 study med start date and start date > end of treatment, assign as post treatment. |
| Known | Partial | Impute stop date as latest possible date:<br>• If only day unknown, impute as the earlier of (last day of the month; date of the last study visit).<br>• If month and day unknown, impute as the earlier of (31st December; date of the last study visit).<br><br>Then:<br>If stop date < 361-203 study med start date, assign as prior.<br>If stop date >= 361-203 study med start date and start date <= end of treatment, assign as concomitant.<br>If stop date >= 361-203 study med start date and start date > end of treatment, assign as post treatment. |
| Known | Missing | If stop date is missing could never be assumed a prior medication.<br>If start date <= end of treatment, assign as concomitant.<br>If start date > end of treatment, assign as post treatment. |
| | | |

| START DATE | STOP DATE | ACTION |
|---|---|---|
| Partial | Known | Impute start date as earliest possible date:<br>*CRF questions: 'Started prior to study?' = Yes; 'Started after last dose of study medication?' = No.*<br>• If only day unknown, impute as the later of (first day of the month; date of birth).<br>• If month and day unknown, impute as the later of (1st January; date of birth).<br>*CRF questions: 'Started prior to study?' = No; 'Started after last dose of study medication?' = Yes.*<br>• If only day unknown, impute as the later of (first day of the month; end of treatment + 1).<br>• If month and day unknown, impute as the later of (1st January; end of treatment + 1).<br>*CRF questions: 'Started prior to study?' = No; 'Started after last dose of study medication?' = No.*<br>• If only day unknown, impute as the later of (first day of the month; date of ICF).<br>• If month and day unknown, impute as the later of (1st January; date of ICF).<br><br>Then:<br>If stop date < 361-203 study med start date, assign as prior.<br>If stop date >= 361-203 study med start date and start date <= end of treatment, assign as concomitant.<br>If stop date >= 361-203 study med start date and start date > end of treatment, assign as post treatment. |

| START DATE | STOP DATE | ACTION |
|---|---|---|
| Partial | Partial | Impute start date as earliest possible date:<br>*CRF questions: 'Started prior to study?' = Yes; 'Started after last dose of study medication?' = No.*<br>• If only day unknown, impute as the later of (first day of the month; date of birth).<br>• If month and day unknown, impute as the later of (1st January; date of birth).<br>*CRF questions: 'Started prior to study?' = No; 'Started after last dose of study medication?' = Yes.*<br>• If only day unknown, impute as the later of (first day of the month; end of treatment + 1).<br>• If month and day unknown, impute as the later of (1st January; end of treatment + 1).<br>*CRF questions: 'Started prior to study?' = No; 'Started after last dose of study medication?' = No.*<br>• If only day unknown, impute as the later of (first day of the month; date of ICF).<br>• If month and day unknown, impute as the later of (1st January; date of ICF).<br><br>Impute stop date as latest possible date:<br>• If only day unknown, impute as the earlier of (last day of the month; date of the last study visit).<br>• If month and day unknown, impute as the earlier of (31st December; date of the last study visit).<br><br>Then:<br>If stop date < 361-203 study med start date, assign as prior.<br>If stop date >= 361-203 study med start date and start date <= end of treatment, assign as concomitant.<br>If stop date >= 361-203 study med start date and start date > end of treatment, assign as post treatment. |

| START DATE | STOP DATE | ACTION |
|---|---|---|
| Partial | Missing | Impute start date as earliest possible date:<br>*CRF questions: 'Started prior to study?' = Yes; 'Started after last dose of study medication?' = No.*<br>• If only day unknown, impute as the later of (first day of the month; date of birth).<br>• If month and day unknown, impute as the later of (1st January; date of birth).<br>*CRF questions: 'Started prior to study?' = No; 'Started after last dose of study medication?' = Yes.*<br>• If only day unknown, impute as the later of (first day of the month; end of treatment + 1).<br>• If month and day unknown, impute as the later of (1st January; end of treatment + 1).<br>*CRF questions: 'Started prior to study?' = No; 'Started after last dose of study medication?' = No.*<br>• If only day unknown, impute as the later of (first day of the month; date of ICF).<br>• If month and day unknown, impute as the later of (1st January; date of ICF).<br><br>Then:<br>If stop date is missing could never be assumed a prior medication.<br>If start date <= end of treatment, assign as concomitant.<br>If start date > end of treatment, assign as post treatment. |
| | | |
| Missing | Known | If stop date < 361-203 study med start date, assign as prior.<br>If stop date >= 361-203 study med start date and CRF question 'Started after last dose of study medication?' = No, assign as concomitant.<br>If CRF question 'Started after last dose of study medication?' = Yes, assign as post treatment. |

| START DATE | STOP DATE | ACTION |
|---|---|---|
| Missing | Partial | Impute stop date as latest possible date:<br>• If only day unknown, impute as the earlier of (last day of the month; date of the last study visit).<br>• If month and day unknown, impute as the earlier of (31st December; date of the last study visit).<br><br>Then:<br>If stop date < 361-203 study med start date, assign as prior.<br>If stop date >= 361-203 study med start date and CRF question 'Started after last dose of study medication?' = No, assign as concomitant.<br>If CRF question 'Started after last dose of study medication?' = Yes, assign as post treatment. |
| Missing | Missing | If CRF question 'Started after last dose of study medication?' = No, assign as concomitant.<br>If CRF question 'Started after last dose of study medication?' = Yes, assign as post treatment. |

## PARTIAL DATE IMPUTATION RULES FOR INITIAL ONSET OF PARKINSON'S DISEASE OR PDP:

For subjects with partial onset dates of Parkinson's Disease or PDP, impute the onset date using the following rules:

- If only day unknown, impute as the earlier of: last day of the month, or date of ICF.
- If both month and day unknown, impute as the earlier of: 31st December of the year, or date of ICF.

# APPENDIX 3. SAS CODE FOR PRIMARY EFFICACY ANALYSIS

| Definition | Variable |
|---|---|
| Subject number (character) | SUBJID |
| Planned double-blind treatment group (numeric) | TRT01PN (1=" SEP-36385 25, 50 or 75 mg/day"; 2="Placebo") |
| Visit (numeric) | AVISITN (4="Week 2"; 5="Week 3"; 6="Week 5"; 7="Week 6") |
| Change from baseline values (numeric) | CHG |
| Baseline values (numeric) | BASE |
| SAPS-PD total score at each post-baseline visit | SAPS-PD 4 to SAPS-PD 7 |

*Mixed Model for Repeated Measures (MMRM)*

```
ODS OUTPUT LSMEstimates=xxxx Coef=xxxx;
PROC MIXED DATA=xxxx;
  CLASS SUBJID TRT01PN AVISITN;
  MODEL CHG = <Baseline SAPS-PD total score> TRT01PN  AVISITN TRT01PN *AVISITN /
DDFM=KR SOLUTION;
  REPEATED AVISITN /SUB=SUBJID TYPE=UN;
  LSMESTIMATE TRT01PN *AVISITN "SEP-363856 at Week 2"  [1, 2 1],
                  "SEP-363856 at Week 3" [1, 2 2],
                  "SEP-363856 at Week 5" [1, 2 3],
                  "SEP-363856 at Week 6" [1, 2 4] / E CL;

  LSMESTIMATE TRT01PN *AVISITN "Placebo at Week 2"  [1, 1 1],
                  "Placebo at Week 3" [1, 1 2],
                  "Placebo at Week 5" [1, 1 3],
                  "Placebo at Week 6" [1, 1 4] / E CL;

  LSMESTIMATE TRT01PN *AVISITN
        "SEP-363856 vs Placebo at Week 2"  [1, 2 1] [-1, 1 1],
```

```
        "SEP-363856 vs Placebo at Week 3" [1, 2 2] [-1, 1 2],
        "SEP-363856 vs Placebo at Week 5" [1, 2 3] [-1, 1 3],
        "SEP-363856 vs Placebo at Week 6" [1, 2 4] [-1, 1 4] / E CL;
RUN;
```

*Pattern-Mixture Model (PMM) with Placebo-Based MI*

Step 1: Partial imputation to get monotone missing pattern:

```
PROC MI DATA=xxxx SEED=12345 NIMPUTE=1000 OUT=xxxx_mono;
  MCMC CHAIN=MULTIPLE INITIAL=EM IMPUTE=MONOTONE NBITER=5000 NITER=200;
  VAR TRT01PN  <Baseline SAPS-PD total score>;
RUN;
```

Step 2: Impute with monotone method:

```
PROC MI DATA=xxxx_mono SEED=12345 NIMPUTE=1 OUT=xxxx_imp;
  VAR  <Baseline SAPS-PD total score>;
  CLASS TRT01PN  ;
    MNAR model(SAPS-PD4/modelobs=(TRT01PN='1'));
    MNAR model(SAPS-PD5/modelobs=(TRT01PN='1'));
    MNAR model(SAPS-PD6/modelobs=(TRT01PN='1'));
    MNAR model(SAPS-PD7/modelobs=(TRT01PN='1'));
  MONOTONE REG;
  BY _Imputation_;
RUN;
```

Step 3: Derive change from baseline at all visits, then fit the primary MMRM on each imputed dataset:

```
ODS OUTPUT LSMEstimates=LSMEST;
PROC MIXED DATA=xxxx_imp;
  CLASS SUBJID TRT01PN  AVISITN;
  MODEL CHG = <Baseline SAPS-PD total score>  TRT01PN  AVISITN TRT01PN *AVISITN /
DDFM=KR SOLUTION COVB;
  REPEATED AVISITN /SUB=SUBJID TYPE=UN;
  LSMESTIMATE TRT01PN *AVISITN "SEP-363856 at Week6"
        [1, 1 4] / E CL;
```

```
  LSMESTIMATE TRT01PN *AVISITN "Placebo at Week 6"
        [1, 2 4] / E CL;
  LSMESTIMATE TRT01PN *AVISITN "SEP-363856 vs Placebo at Week 6"
        [1, 1 4] [-1, 2 4] / E CL;
  BY _Imputation_;
RUN;
```

Step 4: Combine results using Rubin's imputation rules:

```
PROC MIANALYZE parms(classvar=full)=LSMEST;
CLASS TRT01PN *AVISITN;
MODELEFFECTS TRT01PN *AVISITN;
ODS OUTPUT ParameterEstimates=mianLSMEST;
RUN;
```

*Tipping Point Analysis*

Step 1: Impute non-monotone missing data using step 1 from above Pattern Mixture Model (PMM) with Placebo-Based MI.

Step 2: Impute the remaining missing data using a MAR-based regression model, and add a value delta to all imputed values in the SEP-3638528 25, 50 or 75 mg/day group:

```
PROC MI DATA=xxxx_mono seed=12345 NIMPUTE=1 out=xxxx_imp;
  VAR TRT01PN   <Baseline SAPS-PD total score>;
  CLASS TRT01PN  ;
  MONOTONE REG;
    MNAR adjust(SAPS-PD4/delta=xx adjustobs=(TRT01PN ='1'));
    MNAR adjust(SAPS-PD5/delta=xx adjustobs=(TRT01PN ='1'));
    MNAR adjust(SAPS-PD6/delta=xx adjustobs=(TRT01PN ='1'));
    MNAR adjust(SAPS-PD7/delta=xx adjustobs=(TRT01PN ='1'));
  BY _Imputation_;
RUN;
```

Step 3: Repeat steps 3 and 4 from above Pattern Mixture Model (PMM) with Placebo-Based MI to analyze the multiply imputed and shifted (by delta) data, and to combine the results using Rubin's imputation rules.

# APPENDIX 4. STEP 4: REPEAT STEP 2 AND 3 ABOVE FOR ALL VALUES OF DELTA. POTENTIALLY CLINICALLY SIGNIFICANT (PCS) LABORATORY CRITERIA

| Category<br>Parameter Name<br>Age/Gender Restriction, if any | PCS Low | PCS High |
|---|---|---|
| **HEMATOLOGY** | | |
| **WBC** | ≤ 2.8 x $10^9$/L | ≥ 16 x $10^9$/L |
| **Neutrophils (abs)** | < 0.5 x $10^9$/L | > 13.5 x $10^9$/L |
| **Lymphocytes (abs)** | N/A | > 12 x $10^9$/L |
| **Monocytes (abs)** | N/A | > 2.5 x $10^9$/L |
| **Eosinophils (abs)** | N/A | > 1.6 x $10^9$/L |
| **Basophils (abs)** | N/A | > 1.6 x $10^9$/L |
| **Neutrophils (relative)** | ≤ 0.15 | > 0.85 |
| **Lymphocytes (relative)** | N/A | ≥ 0.75 |
| **Monocytes (relative)** | N/A | ≥ 0.15 |
| **Eosinophils (relative)** | N/A | ≥ 0.10 |
| **Basophils (relative)** | N/A | ≥ 0.10 |
| **Hemoglobin** | | |
| Male | ≤ 115 g/L | ≥ 190 g/L |
| Female | ≤ 95 g/L | ≥ 175 g/L |
| **Hematocrit** | | |
| Male | ≤ 0.37 | ≥ 0.60 |
| Female | ≤ 0.32 | ≥ 0.54 |
| **RBC** | ≤ 3.5 x $10^{12}$/L | ≥ 6.4 x $10^{12}$/L |
| **Platelet Count** | ≤ 75 x $10^9$/L | ≥ 700 x $10^9$/L |
| **Ery. Mean Corpuscular Volume (fL)** | ≤ 70 | ≥ 110 |
| **SERUM CHEMISTRY** | | |
| **Sodium** | < 130 mmol/L | > 150 mmol/L |
| **Potassium** | < 3 mmol/L | > 5.5 mmol/L |
| **Chloride** | ≤ 90 mmol/L | ≥ 118 mmol/L |
| **Calcium** | < 1.75 mmol/L | ≥ 3.1 mmol/L |
| **Phosphate** | < 0.65 mmol/L | > 1.65 mmol/L |
| **Bicarbonate** | < 15.1 mmol/L | > 34.9 mmol/L |
| **Magnesium** | < 0.4 mmol/L | > 1.23 mmol/L |
| **AST (IU/L)** | N/A | ≥ 3 x ULN |
| **ALT (IU/L)** | N/A | ≥ 3 x ULN |
| **Alkaline Phosphatase (IU/L)** | N/A | ≥ 1.5 x ULN |
| **CK (IU/L)** | N/A | > 2.5 x ULN |
| **Creatinine** | N/A | ≥ 177 umol/L |
| **BUN** | N/A | ≥ 10.7 mmol/L |

| | | |
|---|---|---|
| **Direct Bilirubin** | N/A | > 2 x ULN |
| **Indirect Bilirubin** | N/A | > 2 x ULN |
| **Bilirubin (mg/dL)** | N/A | ≥ 34.2 umol/L OR > 2 x ULN |
| **Protein** | ≤ 45 g/L | ≥ 100 g/L |
| **Albumin** | ≤ 25 g/L | N/A |
| **Cholesterol** | N/A | > 7.76 mmol/L |
| **HDL-Cholesterol** | < 0.78 mmol/L | N/A |
| **LDL-Cholesterol** | N/A | > 4.14 mmol/L |
| **Triglycerides** | N/A | > 3.42 mmol/L |
| **Uric acid** | | |
| Male | N/A | > 595 umol/L |
| Female | N/A | > 476 umol/L |
| **Glucose** | < 2.78 mmol/L | > 13.9 mmol/L |
| **HbA1c** | N/A | ≥ 0.075 |
| **Prolactin** (Male and Female) | N/A | ≥ 5 x ULN |
| **COAGULATION** | | |
| **aPTT (sec)** | N/A | > 1.5 x ULN |
| **INR (ratio)** | N/A | > 1.5 x ULN |
| **PT (sec)** | N/A | > 1.5 x ULN |
| **URINALYSIS** | | |
| **RBC** | N/A | > 25 hpf |
| **WBC** | N/A | > 25 hpf |

# Document Approval Signature(s)

This document has been signed electronically in compliance with IQVIA and IQVIA Companies' electronic signature policies and procedures. This page is a manifestation of the electronic signatures, which may be confirmed within the IQVIA and IQVIA Companies' source system in accordance with IQVIA and IQVIA Companies' records retention policies. Note: the Title field in the signature block has been intentionally left blank.

UserName: Dufresne, Annie (q103807)
Title:
Date: Tuesday, 25 February 2020, 10:24 AM Eastern Time
Meaning: I Approve the Document
==========

UserName: Woodard, Stacy (q709075)
Title:
Date: Tuesday, 25 February 2020, 10:25 AM Eastern Time
Meaning: I Approve the Document
==========