STATISTICAL ANALYSIS PLAN

SEP380-301

A Multi-region, Multicenter, Randomized, Double-Blind, Placebo-Controlled, Parallel-Group Study Evaluating SEP-4199 Controlled Release (CR) for the Treatment of Major Depressive Episode Associated with Bipolar I Disorder (Bipolar I Depression)

Phase: III

[redacted]

VERSION NUMBER AND DATE: FINAL

DECEMBER 08, 2023

STATISTICAL ANALYSIS PLAN SIGNATURE PAGE

Statistical Analysis Plan Final (Dated 08DEC2023) for Protocol SEP380-301.

		[redacted]	[redacted]	[redacted]
[redacted]		[redacted]		
[redacted]		[redacted]		
[redacted]		[redacted]		

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.

MODIFICATION HISTORY

Unique Identifier for this Version	Date of the Document Version	Author	Significant Changes from Previous Authorized Version
V1.0	08June2023	[redacted]	Not Applicable – First Version
V2.0	03August2023	[redacted]	Incorporated comments on the first draft. Added table of contents for TFL shells.
V2.1	18August2023	[redacted]	Incorporated comments on the draft V2.0. added CV units for lab.
V2.2	10October2023	[redacted]	Simplified the analysis plan because the small sample size due to study early termination.
V2.3	30October2023	[redacted]	Incorporated comments on the draft V2.2.
V2.4	07Novmber2023	[redacted]	Incorporated comments from SMP on the draft V2.3.
Final	08Decmber2023	[redacted]	Incorporated comments on the draft V2.4.

TABLE OF CONTENTS

1. INTRODUCTION

This document describes the rules and conventions to be used in the presentation and analysis of efficacy and safety data for Protocol SEP380-301. It describes the data to be summarized and analyzed, including specifics of the statistical analyses to be performed.

Data and Safety Monitoring Board (DSMB) analysis plan will be described in separate documents. A Blind Data Review (BDR) Plan will be written to describe the process and the outputs to be delivered during the IPD/BDR meetings.

This statistical analysis plan (SAP) is based on protocol version 3.0, dated 06 December 2022. Hereafter, this protocol version is referred to as the Clinical Study Protocol (CSP).

All the final planned analyses specified in this SAP will be performed inhouse by SMPA following the analysis populations, final database lock, and unblinding of treatment.

Some minor modifications may be necessary to the planned design of tables, figures, and listings to accommodate data collected during the actual study conduct.

2. STUDY OBJECTIVES, ESTIMANDS, AND ENDPOINTS

The objective of Study SEP380-301 is to evaluate the efficacy, safety, and tolerability of SEP-4199 CR formulation given as monotherapy at fixed doses of 200 mg/day and 400 mg/day compared with placebo in the treatment of subjects with major depressive episode associated with bipolar I disorder (bipolar I depression).

2.1. PRIMARY EFFICACY OBJECTIVE , ESTIMANDS, AND ENDPOINTS

Evaluate the efficacy of SEP-4199 CR in the reduction of depression symptoms, as measured using the Montgomery-Asberg Depression Rating Scale (MADRS).

Table 1: Primary Objective, Estimand, and Endpoint

Primary	
Objective	**Endpoint**
Evaluate the efficacy of SEP-4199 CR in the reduction of depression symptoms, as measured using the Montgomery-Asberg Depression Rating Scale (MADRS).	Change from baseline to Week 6.in MADRS total score
Estimand	
Population	Subjects with major depressive episode associated with bipolar I disorder, as characterized by the inclusion/exclusion criteria of the study. For the efficacy analyses, the ITT population will be used to represent the population of interest.
Treatments	SEP-4199 CR formulation given as monotherapy at fixed doses of 200 mg/day and 400 mg/day compared with placebo for 6 weeks
Intercurrent events	The intercurrent event that is deemed to have an impact on the interpretation of the variable of interest is early withdrawal from study treatment for any reason.

	This intercurrent event will be handled with the hypothetical strategy. The efficacy data after the last on-treatment visit will not be collected as these data are irrelevant to the treatment effect of interest. Rather, these data will be implicitly predicted based on the assumptions about how the data would evolve after treatment withdrawal.
Population-level summary	The difference in the change from baseline in MADRS total score.

2.2. SECONDARY EFFICACY OBJECTIVE, ESTIMAND, AND ENDPOINT

Evaluate the efficacy of SEP 4199 CR in global improvement of bipolar depression severity, as measured using the Clinical Global Impression-Bipolar Version-Severity of Illness, Depression scale (CGI-BP-S depression).

Table 2: Secondary Objective, Estimand, and Endpoint

Secondary	
Objective	**Endpoint**
Evaluate the efficacy of SEP 4199 CR in global improvement of bipolar depression severity, as measured using the Clinical Global Impression-Bipolar Version-Severity of Illness, Depression scale (CGI-BP-S depression).	Change from baseline to Week 6 in CGI-BP-S depression score.
Estimand	
Population	Subjects with major depressive episode associated with bipolar I disorder as characterized by the inclusion/exclusion criteria of the study. For the efficacy analyses, the ITT population will be used to represent the population of interest.
Treatments	SEP-4199 CR formulation given as monotherapy at fixed doses of 200 mg/day and 400 mg/day compared with placebo for 6 weeks.
Intercurrent events	The intercurrent event that is deemed to have an impact on the interpretation of the variable of interest is early withdrawal from study treatment for any reason. This intercurrent event will be handled with the hypothetical strategy. The efficacy data after the last on-treatment visit will not be collected as these data are irrelevant to the treatment effect of interest. Rather, these data will be implicitly predicted based on the assumptions about how the data would evolve after treatment withdrawal.
Population-level summary	The difference in the change from baseline to Week 6 in CGI-BP-S depression score.

2.3. ADDITIONAL EFFICACY OBJECTIVES AND ENDPOINTS

2.3.1 Additional Efficacy Objectives

- Assess the proportion of subjects with treatment response, defined as ≥ 50% reduction in MADRS total score.

- Assess the proportion of subjects with depression symptom remission, defined as a MADRS total score ≤ 12.
- Assess the effect of treatment on the core symptoms of depression, as measured using the MADRS-6 subscale, defined as summation of items 1, 2, 3, 7, 8, and 9.
- Assess the effect of treatment in the reduction of anxiety symptoms, as measured using the Hamilton Anxiety Rating Scale (HAM-A).
- Assess the effect of treatment on subject self-rated depression symptom severity, as measured by the Quick Inventory of Depressive Symptomatology - Self Report (QIDS-SR16).
- Assess the effect of treatment on functional impairment, as measured by the Sheehan Disability Scale (SDS).
- Assess the effect of treatment on quality of life, as measured by the EuroQol-5 Dimension-5 Level (EQ-5D-5L).
- Assess the effect of treatment on anhedonia, as measured by the Snaith-Hamilton Pleasure Scale (SHAPS).

2.3.2 Additional Efficacy Endpoints

- Change from Baseline in MADRS total score at Weeks 1, 2, and 4
- Change from Baseline in CGI-BP-S depression score at Weeks 1, 2, and 4
- The proportion of subjects with treatment response, defined as ≥ 50% reduction from Baseline in MADRS total score, at Week 6
- The proportion of subjects meeting criteria for remission, defined as MADRS total score ≤ 12, at Week 6
- Change from Baseline to Week 6 in the MADRS-6 core symptoms subscale
- Change from Baseline in HAM-A total score at Weeks 1, 2, 4, and 6
- Change from Baseline to Week 6 in the QIDS-SR16 total score
- Change from Baseline to Week 6 in the SDS total score and subscale scores (Work/School, Family, and Social function)
- The proportion of subjects meeting criteria for functional remission, defined as having a score ≤ 2 on each of the SDS subscale scores (Work/School, Family, and Social function) at Week 6
- Change from Baseline to Week 6 in the EQ-5D-5L Index score and VAS score
- Change from Baseline to Week 6 in the SHAPS total score

2.4. SAFETY OBJECTIVES AND ENDPOINTS

2.4.1 Safety Objectives

- Determine the incidence of adverse events (AEs), discontinuation due to AEs, serious AEs (SAEs), and adverse events of special interest (AESI).
- Evaluate safety and tolerability using physical examinations, 12-lead electrocardiograms (ECG), vital signs, clinical laboratory tests, prolactin levels, metabolic parameters, body weight, and body mass index (BMI).
- Monitor for akathisia and extrapyramidal symptoms (EPS) using the Abnormal Involuntary Movement Scale (AIMS), Barnes Akathisia Scale (BARS), and Modified Simpson-Angus Scale (SAS).
- Monitor for treatment-emergent mania or hypomania, defined as a Young Mania Rating Scale (YMRS) score of ≥ 16 on any 2 consecutive visits or at the final assessment, or an AE of mania or hypomania.
- Monitor for suicidality using the Columbia-Suicide Severity Rating Scale (C-SSRS).
- Evaluate potential withdrawal symptoms using the Physician's Withdrawal Checklist (PWC).

2.4.2 Safety Endpoints

- The incidence of overall AEs, discontinuation due to AEs, and SAEs
- The incidence of hyperprolactinemia-related AESI by sex and overall
- Clinical laboratory evaluations (serum chemistry, hematology, thyroid panel, and urinalysis)
- Clinical evaluations (vital signs including orthostatic effects, and 12-lead ECGs)
- Change and percent change from Baseline to Week 6 in body weight (kg)
- Change from Baseline to Week 6 in BMI (kg/m^2)
- Changes from Baseline in metabolic parameters, including glucose, hemoglobin A1c (HbA1c), insulin, Homeostatic Model Assessment for Insulin (HOMA-IR), total cholesterol, high density lipoprotein (HDL), low density lipoprotein (LDL), and triglycerides at Week 6
- Incidence of treatment-emergent mania, defined as a YMRS total score ≥ 16 at two consecutive visits or at final visit, or an AE of hypomania or mania
- Mean changes from Baseline and the proportion of subjects with worsening on the movement disorders scales: AIMS, BARS, and modified SAS at Week 6
- Frequency and severity of suicidal ideation and suicidal behavior as assessed by the C-SSRS
- Occurrence of potential symptoms of withdrawal from SEP-4199 CR measured by change from Week 6 in the PWC total score at the safety Follow-up Visit

2.5. PHARMACOKINETIC AND PHARMACODYNAMIC OBJECTIVES AND ENDPOINTS

2.5.1 Pharmacokinetic and Pharmacodynamic Objectives

- Evaluate the therapeutic plasma concentration range of SEP-4199 CR taken as 200 mg/day and 400 mg/day for treatment of major depressive episode associated with bipolar I disorder.
- Perform population pharmacokinetic (Pop-PK) analysis using plasma concentrations of SEP-4199 CR 200 mg/day and 400 mg/day.
- Evaluate the relationship between SEP-4199 PK and plasma prolactin levels for SEP-4199 CR 200 mg/day and 400 mg/day.
- Explore the exposure-response relationship of SEP-4199 CR 200 mg/day and 400 mg/day and symptoms as measured by MADRS using population pharmacokinetic (PK)/ pharmacodynamic (PD) methods.

2.5.2 Pharmacokinetic and Pharmacodynamic Endpoints

- Plasma concentrations of aramisulpride and esamisulpride
- Plasma concentrations of prolactin

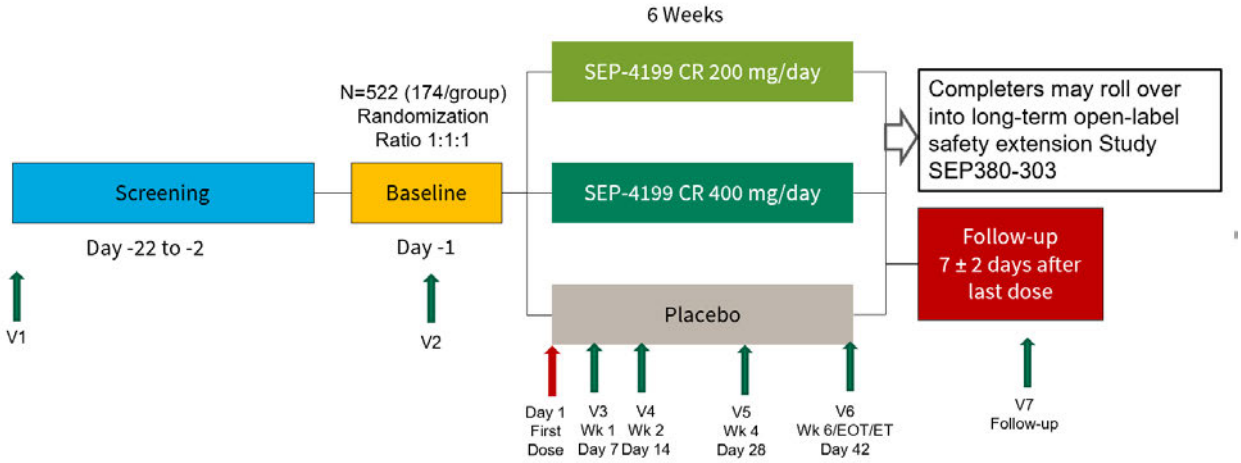
3. STUDY DESIGN

3.1. GENERAL DESCRIPTION

SEP380-301 is a randomized, double-blind, placebo-controlled, parallel-group study designed to evaluate the efficacy, safety, and tolerability of treatment with SEP 4199 CR at fixed doses of 200 mg/day or 400 mg/day compared with placebo for the treatment of major depressive episode associated with bipolar I disorder (bipolar I depression). The study is projected to randomize approximately 522 subjects in North America, Latin America, Japan, and Europe to SEP 4199 CR 200 mg/day, SEP 4199 CR 400 mg/day, and placebo treatment groups in a 1:1:1 ratio, resulting in approximately 174 subjects/group.

The study will consist of a Screening Period (up to 21 days), a 6 week double-blind Treatment Period (42 days), and a Follow-up Period (7 [± 2] days after the last study drug dose), as shown in the following figure. If necessary, subjects may return to the clinic at any time for an unscheduled visit. Subjects who complete the Treatment Period are eligible to enroll directly into a long-term open-label safety extension study (SEP380-303) of SEP 4199 CR. Those subjects who prematurely discontinue or who complete the Treatment Period and choose not to enroll in the long-term open-label safety extension study will have a follow-up safety Visit 7 (± 2) days after their last dose of double-blind study drug.

Figure 1 Study Schematic



Abbreviations: CR = controlled release; EOT = end of treatment; V = visit; Wk = week

Note: Subjects who do not enroll in long-term open-label safety extension study will have Visit 7. For Japan, there is an extra visit on Day 1.

3.2. METHOD OF ASSIGNING SUBJECTS TO TREATMENT GROUPS

Subjects who successfully meet eligibility criteria will be randomly assigned to receive SEP 4199 CR 200 mg/day, SEP 4199 CR 400 mg/day, or matching placebo in a double blind fashion in a 1:1:1 ratio. The randomization will be balanced using permuted blocks with stratification factor of countries. The stratification process will be handled in an Interactive Web Response Systems (IXRS).

At the Baseline Visit (Visit 2), the treatment code, which is linked to the randomization schedule, will be assigned by the IXRS designated by the Sponsor. A unique subject number will be assigned by the IXRS when a subject enters the Screening Period. Each subject will be given one subject number comprised of 9 digits. Each subject number will specify the study (3 digits), site (3 digits), and subject (3 digits) (eg, 301001001 would denote: Study 301, Site 001, and Subject 001). If a subject does not meet study entry criteria, his or her subject number cannot be reassigned to another subject.

Subjects may be screened up to a maximum of 3 times, if judged appropriate by the Investigator and approved by the Medical Monitor. Each time the subject is rescreened, they will be reconsented and receive a new subject number, which cannot be reassigned to another subject. Subjects who do not pass Inclusion Criterion#12 (ie, subject meets an additional inclusion criterion at Baseline that will remain blinded to clinical site Investigators and staff) may not be re-screened.

3.3. BLINDING

Subjects, Investigators, clinical site staff, persons performing the assessments, clinical operations personnel (including the Sponsor's bioanalytical manager), data analysts, personnel at central laboratories (including imaging), and the Sponsor will remain blinded to randomization scheme until blinding is formally broken for all subjects. For this to occur, all subjects must have completed the study and the study database must be locked.

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 following exceptions: bioanalytical laboratory personnel involved in the analysis of PK samples, Data and Safety Monitoring Board (DSMB) members involved in regular review of safety data, external statistical staff involved in preparing materials for DSMB reviews, and the Sponsor's clinical trials materials management.

Prolactin levels will be masked in regular lab data transfer for any assessment collected after the first study drug dose and until database is locked and treatment is unblinded (that is to say, prolactin assessed prior to the first study drug dose will be unmasked and available for study personnel during the study).

The identity of the treatments will be concealed by the use of study drugs that are all identical in packaging, labeling, schedule of administration, and appearance.

Plasma concentrations of aramisulpride, esamisulpride, and total amisulpride will not be disclosed before unblinding. In any case that concentration data transfer to the Sponsor becomes necessary prior to the database lock, then the concentration data can only be shared with the Sponsor's bioanalytical project manager with dummy subject IDs.

3.4. DETERMINATION OF SAMPLE SIZE

The sample size is estimated based on the primary efficacy endpoint, the change from Baseline in MADRS total score at Week 6. Based on the Phase 2 Study SEP380-201 results for both planned and post-hoc analyses of the primary endpoint, assuming a common standard deviation (SD) of 10 and a mean improvement of 3.5 (effect size 0.35) and 3.5 (effect size 0.35) over placebo for SEP 4199 CR 200 mg and 400 mg dose groups, respectively, with a randomization ratio of 1:1:1 for placebo, SEP 4199 CR 200 mg and 400 mg, a sample size of 165 subjects per treatment group will provide at least 80% conjunctive power to reject both null hypotheses of no difference between SEP 4199 CR doses and placebo and at least 90% power to reject at least one individual null hypothesis after multiplicity adjustment for two comparisons of the primary endpoint using the enhanced mixture truncated Hochberg gatekeeping procedure (Kordzakhia 2018) with truncation parameter 0.5 and 2-sided Type I error of 0.05. An upward adjustment of approximately 5% is assumed to compensate for subjects who are randomized but have no post-Baseline efficacy assessment. Thus, a total sample of approximately 174 subjects per group (522 subjects in total) will be randomized. The sample size calculation is based on 100,000 Monte Carlo computer simulations.

3.5. CHANGES IN THE CONDUCT OF THE STUDY

Effective 10-Oct-2023 the Sponsor has decided to discontinue SEP380-301 & SEP380-303 studies early based on numerous considerations including operational feasibility of conducting and completing the studies in a timely manner. There were no emergent safety, tolerability or efficacy issues associated with the investigational study medication (SEP-4199 CR) that contributed to this decision.

3.6. SCHEDULE OF EVENTS

Schedule of events can be found in Section 1, Table 2 Schedule of Assessments of the CSP. This table is also included in APPENDIX 14 of the SAP. There is an additional visit for Japan on Day1 which checks study drug compliance and ECG.

3.7. CHANGES TO ANALYSIS FROM PROTOCOL

Because of the small sample size due to study early termination, the following main changes will be made for this study:

1. Except for disposition, demographics, some key efficacy and safety variables, combined SEP-4199 CR vs. placebo will be performed for the efficacy analyses using ITT population and safety analyses using Safety population.
2. Only include Safety population, ITT population, and PK population, no analyses using PP population and Follow-up population will be performed.
3. Except for PWC, follow-up period analysis will not be performed.
4. No subgroup analyses and sensitivity analyses will be performed.
5. No multiple comparisons for the primary and secondary efficacy endpoints will be performed.
6. SDS subscale scores and proportion of subjects meeting functional remission criteria per the SDS subscales will not be analyzed.
7. AE will not be summarized by high-level term (HLT).
8. Rank ANCOVA analysis will not performed for selected lab, vital signs, and ECG parameters.
9. To be consistent with AIMS and BARS, MSAS will be analyzed by total score.

4. PLANNED INTERIM ANALYSIS AND SAFETY MONITORING ANALYSIS

The following interim analysis and safety monitoring analysis will be performed for this study.

4.1. DATA AND SAFETY MONITORING BOARD (DSMB)

A Data and Safety Monitoring Board (DSMB) will review safety data at regular intervals. The DSMB will be independent of the Sponsor, CRO, and the Investigators and will be empowered to recommend stopping the study due to safety concerns, but not for efficacy or futility. The DSMB may review blinded, unblinded, or partially unblinded data, but the Sponsor (with the exception of the relevant members of the pharmacovigilance team responsible for reporting Suspected Unexpected Serious Adverse Reactions [SUSARs]), the CRO, and the Investigators will remain blinded until the official unblinding of the database. The membership of the DSMB and its mandate will be described in a separate DSMB charter.

4.2. INTERIM ANALYSIS

Other than the DSMB safety reviews, there is no interim analysis planned for this study.

5. ANALYSIS POPULATIONS

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

5.1. INTENTION-TO-TREAT [ITT] POPULATION

The intention-to-treat (ITT) population will consist of subjects who are randomized and received at least one dose of study medication and have both baseline and at least one post-Baseline assessment in either MADRS or CGI-BP-S depression scores. The ITT population is the primary population for efficacy analyses, ie, unless otherwise specified, all efficacy analyses will be based on the ITT population. Subjects in the ITT population will be analyzed based on the treatment to which they are randomized.

5.2. SAFETY [SAF] POPULATION

The safety population will consist of all subjects who randomized and receive at least one dose of study medication. Subjects will be analyzed according to modal dose (the predominant treatment) received. The modal dose or predominant treatment is defined as the treatment to which the subject is exposed for the greatest duration during the treatment period. This will generally be the same as the randomized treatment group, unless the subject takes incorrect study medication during their entire participation in the study.

5.3. PHARMACOKINETICS (PK) POPULATION

The PK population includes all subjects who are randomized, receive at least 1 dose of study drug, and have any post-Baseline PK concentrations of SEP-4199 CR. PK samples collected from placebo subjects will not be analyzed.

The PK analysis will be based on the PK population.

Population PK (Pop-PK) analysis methods will be used to characterize the PK/PD profiles in subjects treated with SEP 4199. These methods will be described in a separate document and will not be included in the SAP.

6. GENERAL CONSIDERATIONS

6.1. REFERENCE START DATE AND STUDY DAY

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

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

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

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

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

Study Day = (date of assessment or event – reference date).

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

6.2. BASELINE

Unless otherwise specified, baseline is defined as the last non-missing value taken prior to the first dose of study medication (including unscheduled assessments).

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 the last non-missing value and the reference start date coincide, that value will be considered the baseline. Adverse Events and medications commencing on the reference start date will be considered postbaseline.

6.3. DERIVED TIME POINTS

As per protocol, Visit 6 can be an End of Treatment (EOT) visit or an Early Termination (ET) visit. If a subject terminates early, his/her ET visit will be mapped to the next planned assessment visit if it is within the window of the next planned assessment visit and within 7 days after treatment discontinuation. This applies to both efficacy and safety data. If assessments from an ET visit cannot be mapped to the next planned assessment visit, these assessments will be excluded from any mixed model for repeated measures (MMRM). No follow-up visit data will be included in any MMRM/ANCOVA analyses. In the case of the ET being out of window of the next planned visit, it will be mapped to an unscheduled visit as appropriate.

All applicable by-visit efficacy and safety tables will include descriptive statistics for the Follow-up visit and the "Last Observation Carried Forward (LOCF) visit". The "Last Observation Carried Forward (LOCF) visit" is defined as the post-baseline visit at which the last non-missing observation is collected for a subject for a given parameter, including unscheduled visits or unmapped ET visits, but excluding follow-up visit. The "Last Observation Carried Forward (LOCF) visit" will be excluded from any MMRM analyses.

For remapping of the ET visits, see Table 3 below.

Table 3: Mapping of the ET visit.

Parameter	Early Termination Day Criteria	Week	Analysis Visit
ECG, Vital Sign, YMRS, C-SSRS, Concomitant Medication, Last dose of Benzodiazepines/ Sedatives/Hypnotics	1 ≤ study day ≤ 9	1	Week 1
	10 ≤ study day ≤ 16	2	Week 2
	17 ≤ study day ≤ 30	4	Week 4
	31 ≤ study day ≤ 44	6	Week 6
MADRS, CGI-BP-S, HAM-A	1 ≤ study day ≤ 9	1	Week 1
	10 ≤ study day ≤ 16	2	Week 2
	17 ≤ study day ≤ 30	4	Week 4
	31 ≤ study day ≤ 44	6	Week 6
QIDS-SR16, SDS, EQ-5D-5L, SHAPS, AIMS, BARS, MSAS Weight, BMI, Waist Circumference, PWC, HbA1C, Lipid panel, Serum Insulin, HOMA-IR, CRP	31 ≤ study day ≤ 44	6	Week 6
Hematology, Serum Chemistry*, Urinalysis, Serum Prolactin, PK, Plasma Prolactin	10 ≤ study day ≤ 16	2	Week 2
	31 ≤ study day ≤ 44	6	Week 6

* Except for parameters listed above such as HbA1C, Lipid panel, Serum Insulin, HOMA-IR, CRP.
Abbreviations: AIMS = Abnormal Involuntary Movement Scale; BARS = Barnes Akathisia Rating Scale; BMI = Body Mass Index; CGI-BP-S = Clinical Global Impression Bipolar Version, Severity of Illness; C-SSRS = Columbia-Suicide Severity Rating Scale; ECG = 12-Lead Electrocardiogram; ET = Early Termination; EQ-5D = EuroQoL-5D; HAM-A = Hamilton Rating Scale for Anxiety; MADRS = Montgomery-Asberg Depression Rating Scale; PWC = Physician's Withdrawal Checklist; QIDS-SR16 = Quick Inventory of Depressive Symptomatology – Self-Report; SAS = Modified Simpson-Angus Scale; SDS = Sheehan Disability Scale; YMRS = Young Mania Rating Scale.

All planned visits will be mapped to weeks and analysis visits for summaries and statistical analyses where applicable (Table 4).

Table 4: Mapping of planned visits.

Planned Visit	Analysis Week	Analysis Visit
Visit 1	-1	Screening
Visit 2	0	Baseline
Visit 2 (day 1 Japan)		Day 1
Visit 3	1	Week 1
Visit 4	2	Week 2
Telephone Contact 1 (TC1)	3	TC1
Visit 5	4	Week 4
Telephone Contact 2 (TC2)	5	TC2
Visit 6	6	Week 6
Visit 7	7	Follow-up

Original visit collected on the case report forms (CRFs) will be displayed in the listings.

6.4. WINDOWING CONVENTIONS

Apart from early termination data (see Section 6.3), no visit windowing will be performed during the analysis for this study. Data will be summarized by analysis visit and analyzed by week where applicable (see Section 6.3).

6.5. RETESTS, UNSCHEDULED VISITS AND EARLY TERMINATION DATA

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

Unscheduled assessments will not be included in by-visit summaries. Unscheduled assessments collected prior to the first dose of study drug will contribute to the derivation of the baseline value. Unscheduled assessments collected postbaseline will contribute to the derivation of LOCF endpoint, potentially clinically significant (PCS) postbaseline value, and best/worst case value where required (eg, shift tables).

In the case of a retest, the assessment recorded under the scheduled visit will be used for by-visit summaries, and the assessment(s) recorded under unscheduled visit(s) will be presented in listings only.

Early termination data collected postbaseline will be assigned to the next planned visit for that assessment as explained in Section 6.3. This mapping will be implemented to all early termination data

points used in the efficacy and safety analyses.

Listings will include scheduled, unscheduled (including retests), and early termination data with original dates and visits displayed.

6.6. STATISTICAL TESTS

All statistical inference will be performed with 2-sided tests at the significance level of 0.05 and 2-sided 95% confidence intervals (CIs) whenever appropriate. All data will be summarized by treatment group and visit as appropriate. All subject data will be presented in data listings by subject.

6.7. COMMON CALCULATIONS

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

- Assessment Value at Visit X – Baseline Value

and percentage change from baseline will be calculated as:

- 100 (Assessment Value at Visit X – Baseline Value)/Baseline Value.

6.8. SOFTWARE VERSION

All analyses will be conducted using Statistical Analysis System (SAS®) version 9.4 or higher (SAS Institute, Inc., Cary, North Carolina).

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 centers in the Europe, Japan, North America, and Latin America.

When specified, statistical analysis will be adjusted for region with centers within the region pooled. Region will be categorized as follows:

- USA
- Europe

- Japan

7.3. MISSING DATA

For the MMRM models, missing observations are treated as missing at random (MAR) and no imputation for missing data will be applied.

Unless otherwise specified, any individual missing item in any scale will not be imputed. When calculating a total score, subscale score, or any assessment with more than one item, if one or more items are missing at a visit, then the associated score (ie, total score or subscale score) will be set to missing. For additional details, see the individual scale description sections.

Handling of missing efficacy data, if any, is described in Sections 16.1.2, 16.2.2, 16.3.2.

Handling of missing safety data, if any, is described in Sections 17.1.1.1, 17.1.1.2, and 17.6.

Handling of missing PK data if any is described in Section 18.

See APPENDIX 2 for details of incomplete/missing dates.

8. OUTPUT PRESENTATIONS

APPENDIX 1 shows conventions for presentation of data in outputs.

The shells 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 in separate documents.

9. DISPOSITION AND WITHDRAWALS

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

Subject disposition will be presented for the screened subjects described by the randomized treatment group (where applicable) and overall for all subjects. The number and percentage of subjects who were screened, screen-failed (with reasons for screen failure), randomized, randomized but not dosed, received study medication, completed or discontinued (with reasons for discontinuation) from the study treatment period, and completed the follow-up visit, and completed the treatment period but not the follow-up visit will be presented. The number of subjects with non-missing change from baseline at Week 6 of MADRS total score and the number of subjects with non-missing change from baseline at Week 6 of CGI-BP-S depression score will also be presented.

With respect to the above, the following definitions apply:

- Screened Subjects: Any subject who signed the study specific informed consent and completed at least 1 study-related procedure.
- Screen failure: Subjects who signed the study specific informed consent but either failed to meet screening requirements or was successfully screened but was not randomized.
- Randomized Subjects: Any subject who was randomized into the treatment period of the study and was assigned a randomization number.

The number of subjects randomized will also be described by site, country, and region by treatment group and overall. The number of subjects who complete (including remotely versus in-person, for

Week 6 and the Follow-up visit only) and discontinue from the study medication before or at a given visit, along with reasons for study medication discontinuation will be described by the treatment group and overall using randomized subjects.

The number of subjects in each population will also be described.

10. IMPORTANT PROTOCOL DEVIATIONS

A complete list of subjects who meet at least one of important protocol deviations (IPDs) criteria, will be identified through a blinded data review meeting (BDRM) prior to database lock. Details of the IPD categories and criteria are provided in the IPD specifications document. IPDs will be identified for all randomized subjects and presented in a data listing.

In addition, the number and percentage of subjects within each IPD category will be summarized for the SAF population by the actual treatment group.

10.1. DEVIATIONS RELATED TO PK ANALYSIS

Changes to the procedures or events, which may impact the quality of the PK data, may be considered significant protocol deviations and will be described within the clinical study report body text. These changes or events will include any circumstances that will alter the evaluation of the PK. Examples include, but may not be limited to, sample processing errors that lead to inaccurate bioanalytical results, and/or inaccurate dosing on the day of PK sampling. In the case of a significant protocol deviation or event, PK data collected for the affected dose group will be excluded from the study summary and inferential results but listed, if appropriate. Other changes to the procedures or events which do not impact the quality of the PK data will not be considered significant. A common example of a non-significant protocol deviation is a missed blood sample or deviations from blood collection times.

11. DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

Demographic data and other baseline characteristics will be presented for the SAF population. The data will be presented by treatment group (SEP-4199 CR 200 mg/day, SEP-4199 CR 400 mg/day, all SEP-4199 CR, and Placebo) and overall by the predominant treatment received.

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

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 years
    - ≥ 55 years
- Gender
- Race
    - American Indian or Alaska Native

  - Asian
  - Black or African American
  - Native Hawaiian or Other Pacific Islander
  - White
  - Other
- Ethnicity
  - Hispanic or Latino
  - Not Hispanic or Latino
- Country
  - Bulgaria
  - Romania
  - USA
  - Japan
- Region
  - USA
  - Europe
  - Japan
- Weight (kg)
- Height (cm)
- 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 MADRS total score, as a continuous variable
- Baseline CGI-BP-S depression score, as a continuous variable

Demographic data collected for all screened subjects will be listed.

The following psychiatric history data will be summarized for the SAF population in a separate table:

- Age at onset of bipolar I disorder
- Time since initial onset of bipolar I disorder, in years, calculated relative to date of informed consent (see APPENDIX 2 for partial date imputation rules)
- Time since onset of current episode of major depression associated with bipolar I disorder symptoms, in months, calculated relative to date of informed consent (see APPENDIX 2 for partial date imputation rules)
- Frequency distribution of number of prior hospitalizations for Bipolar I depression (0, 1, 2, 3, and 4 or more) to date of informed consent (see APPENDIX 2 for partial date imputation rules).

- Frequency distribution of DSM-5 code in ascending numeric order
- Bipolar I diagnosis subtype (Bipolar I disorder without rapid Cycling and Bipolar I disorder with rapid cycling)
- Number and percentage of subjects with any other psychiatric disorders

Summaries of subjects' family psychiatric history will also be presented for the Safety population.

Other psychiatric Disorder data will be summarized for SAF population in a separate table by summarizing Diagnosis and DSM-5 codes.

11.1. DERIVATIONS

- Time since initial onset of bipolar disorder (years):

  (Date of informed consent – date of initial onset of bipolar disorder) / 365.25

- Time since initial onset of current episode of major depression (months):

  (Date of informed consent – date of initial onset of current episode of major depression) / 30.4375

12. MEDICAL HISTORY

Medical History information will be presented for the SAF populations by treatment group (all SEP-4199 CR and Placebo) and overall. The data will be presented by the predominant treatment received. 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 MedDRA, Version 24.1 or higher, and presented by SOC and PT. Medical History will be sorted alphabetically by SOC and by decreasing frequency of PT in the All SEP-4199 CR column.

Summaries of subjects' family medical history will be presented for the Safety population.

Reproductive history of females will also be summarized for SAF population in a separate table.

13. PRIOR/CONCOMITANT/POST-TREATMENT MEDICATIONS

Medications will be coded to Anatomical Therapeutic Chemical (ATC) levels and Preferred Names using World Health Organization Drug Dictionary Global, Version SEP 2021. Medications will be presented for the SAF population and coded to indication-specific Anatomical Therapeutic Chemical (ATC) Level 3 and Preferred Name. If ATC level 3 is missing, then ATC level 2 will be used. Medication will be sorted alphabetically by ATC level 3 and by decreasing frequency of Preferred Name in the All SEP-4199 CR column.

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; ie, prior, concomitant, and post-treatment.

- Prior medications are medications which started prior to the first dose of study medication.

- Concomitant medications are medications that:
    - Started prior to, at the time of, or after the first dose of study medication, AND
    - Started before or at the date (day) of the last dose of study drug, AND
    - Ended at the time of or after the first dose of study drug or were ongoing at the end of the study.
- Post-treatment medications are medications which ended after the day of the last dose of study drug or ongoing.

Prior and concomitant medication use will be summarized by treatment group (all SEP-4199 CR and Placebo) by ATC Level 3 classification and Preferred Name using the SAF population. Subjects with multiple uses of a medication will be counted only once for a given ATC level 3 class or Preferred Name. Number and percentage of subjects who take concomitant benzodiazepines, sedatives, or hypnotics will be summarized by visit (including follow-up visit) and overall. Prior, concomitant, and post-treatment medications will be provided in data listings.

14. STUDY DRUG EXPOSURE

Exposure to study drug will be calculated as the number of days from first dose date to last dose date. Duration of exposure to study drug will be summarized for the SAF population.

The date of first study drug administration will be taken from the CRF “Study Drug Administration / Drug Accountability” form.

The date of last study drug will be retrieved as the latest date among “Study Drug Administration / Drug Accountability” and “Subject Disposition - End of Study” forms. For subjects who are lost to follow-up, and for whom the last dose date is unknown, the last contact date will be used.

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

Overall number of tablets taken will be summarized by treatment group (SEP-4199 CR 200 mg/day, SEP-4199 CR 400 mg/day, all SEP-4199 CR, and Placebo) as a continuous variable.

Duration of exposure to study medication in days will be summarized both as a continuous variable and categorically, and by planned visit days:

- Number and percentage of subjects with drug exposure ≥ 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.

Total person years of exposure will be calculated by treatment group (SEP-4199 CR 200 mg/day, SEP-4199 CR 400 mg/day, all SEP-4199 CR, and Placebo).

14.1. DERIVATIONS

- Duration of exposure (days)

  Last dose of study drug – First dose of study drug + 1.

- Person years of exposure

Sum of total exposure (days) for all subjects within each treatment group / 365.25

- Overall number of tablets taken

  Sum of number of dispensed tablets – Sum of number of returned tablets, if any returned tablets is missing then overall number of tablets taken is set to missing

15. STUDY DRUG COMPLIANCE

Compliance to study drug will be presented by treatment group (all SEP-4199 CR and Placebo) for the SAF population.

Percent compliance will be calculated by visit and overall on the number of subjects with non-missing compliance data. For each postbaseline visit, if the number of tablets returned is missing, the corresponding compliance will be set as missing. Overall compliance will be set to missing if one or more blister cards are not returned to the site for accountability.

Non-compliance is defined as less than 75% or more than 125% non-missing compliance. Subjects with missing compliance will not be classified as non-compliant but described in the categorical analysis.

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

Subject level treatment compliance will be listed.

15.1. DERIVATIONS

Compliance with study drug will be calculated for each period defined by study visits (ie, per-visit compliance) and overall.

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

$$\frac{\#\text{ Tablets dispensed at Visit (V-1)} - \#\text{ Tablets returned at Visit V}}{\#\text{ Tablets should be taken per day} \times (\text{Date of Visit V} - \text{Date of Visit (V-1)})} \times 100\%$$

For week 1, the date of first dose of study medication will be used instead of date of visit (V-1). Two tablets per day are supposed to be taken. If number tablets dispensed at Visit (V-1) and/or number tablets returned at Visit V are missing, per-visit compliance will not be calculated for the period impacted.

If a subject discontinued from the study in between Visit (V-1) and Visit V except for subjects lost to follow-up in which case ET visit was not done, then the above formula will be modified to:

$$\frac{\#\text{ Tablets dispensed at Visit (V-1)} - \#\text{ Tablets returned at Visit ET}}{\#\text{ Tablets should be taken per day} \times (\text{Date of Visit ET} - \text{Date of Visit (V-1)})} \times 100\%$$

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:

$$\frac{\text{Total \# tablets dispensed} - \text{Total \# tablets returned}}{\#\text{ Tablets should be taken per day} \times \text{Duration of Exposure}} \times 100\%$$

Duration of exposure is calculated as specified in Section 14.1.

If number tablets dispensed and/or number tablets returned are missing at 1 or more visits, overall compliance will not be calculated.

16. EFFICACY OUTCOMES

16.1. PRIMARY EFFICACY

16.1.1. PRIMARY EFFICACY VARIABLE & DERIVATION

The primary efficacy endpoint is the change from baseline in the MADRS total score at Week 6.

The MADRS is a clinician-rated assessment of the subject's level of depression. The measure contains 10 items that measure apparent and reported sadness, inner tension, reduced sleep and appetite, difficulty concentrating, lassitude, inability to feel, and pessimistic and suicidal thoughts. Each item is scored in a range of 0 to 6 points, with higher scores indicating increased depressive symptoms. Total score will be equal to the sum of the 10 items (range between 0 and 60). MADRS is assessed at Visit 1/Screening, Visit 2/Baseline/Week 0, Visit 3/Week 1, Visit 4/Week 2, Visit 5/Week 4, and Visit 6/Week 6.

16.1.2. MISSING DATA METHODS FOR PRIMARY EFFICACY VARIABLE

The MADRS total score at a visit will be set to missing at that visit if any one item is missing. The primary efficacy variable, change from baseline in MADRS total score at Week 6, will be set to missing if MADRS score at Week 6 is missing or baseline value is missing. The same applies to all other visits.

The primary analysis of the primary efficacy variable will use an MMRM which makes an MAR assumption to the missing MADRS total score. Early termination data will be mapped as described in Section 6.3. Missing data will not be imputed.

16.1.3. PRIMARY ANALYSIS OF PRIMARY EFFICACY VARIABLE(S)

The primary efficacy estimand is defined as the efficacy of all SEP-4199 CR over placebo for 6 weeks, calculated as the difference in the primary efficacy endpoint between all SEP-4199 CR treatment group and the placebo treatment group in the ITT population. This estimand provides an estimate of the efficacy of all SEP-4199 CR over placebo, should a subject be able to tolerate and adhere to treatment for 6 weeks.

The 4 attributes of the primary efficacy estimand are as follows:

A. **Population of interest:** ITT subjects with major depressive episode associated with bipolar I disease (bipolar I disorder) defined through the inclusion/exclusion criteria.

B. **The variable of interest (endpoint):** change from baseline in MADRS total score at Week 6

C. Intercurrent events (and how to handle them)

The estimand is based on the effect of treatment in the hypothetical absence of any intercurrent events, ie, if all subjects had adhered to treatment up to Week 6. This includes events where a subject continues the study regardless of the intercurrent event(s) (eg, subject received a prohibited medication during the treatment period and was not discontinued from the study) and/or events that cause the

subject to discontinue from the study.

Any intercurrent event where subject continues the study regardless of the intercurrent event will be ignored for the primary analysis and the primary efficacy data will be analyzed as is.

The statistical analysis for intercurrent events that lead to treatment discontinuation rests on an assumption about assessments that would have been observed under the hypothetical setting where a subject does not discontinue from treatment. For example, a subject who discontinues treatment before Week 6 in the all SEP-4199 CR treatment group is assumed to behave similar to a subject who completes the study up to Week 6 in the same SEP-4199 CR treatment group using the MMRM analysis based on the MAR assumption. The assumption for subjects in the placebo group is made similarly.

D. Population level summary of the variable:

Difference in means of change from baseline in MADRS total score at Week 6 comparing all SEP-4199 CR treatment group to placebo

The efficacy of SEP-4199 CR in terms of the MADRS total score will be evaluated using the following null hypothesis:

H_1: There is no difference in mean change from baseline at Week 6 on the MADRS total score in the all SEP-4199 CR treatment group compared to Placebo.

The alternative hypothesis for the null hypothesis is that there is a difference.

The primary efficacy endpoint will be analyzed for the ITT population using an MMRM analysis, with change from baseline in the MADRS total score as the response variable and with factors for treatment group, week, region, baseline MADRS total score, and the treatment-by-week interaction. This analysis method makes an MAR assumption to the missing primary efficacy endpoints. The Kenward Roger approximation will be used to estimate denominator degrees of freedom. An unstructured covariance matrix will be used for the within-subject correlation. Heterogeneous Toeplitz and Toeplitz, and Compound Symmetry will be assumed sequentially in case the model fails to convergence. The first covariance structure to yield convergence will be used in the analysis. For the primary efficacy endpoint, two models are used: one for all SEP-4199 CR versus placebo, one for each SEP-4199 CR dose group versus placebo group.

The Least Square (LS) mean treatment difference of change from baseline at Week 6, their 2-sided 95% CIs, and the associated p-value will be calculated based on the MMRM.

At each postbaseline visit, effect-size will be presented. Based on the MMRM, effect size at a visit will be calculated as the absolute value of the LS means difference from placebo divided by the model estimate of the pooled SD at the visit, which is obtained from the square root of the diagonal element, associated at the visit, from the covariance matrix (R matrix of subjects with MADRS total score at all visits).

16.1.4. Supportive Analysis of Primary Efficacy Variable(s)

The primary efficacy endpoint will be analyzed using an analysis of covariance (ANCOVA) on the ITT population as supportive analysis for each scheduled postbaseline visit and also Week 6 LOCF endpoint. The model will include terms for treatment, region, and baseline MADRS total score as covariate. The LS mean of treatment difference (all SEP-4199 CR group minus placebo), the 2-sided 95% CIs, and the associated p-values will be obtained from the model.

Based on the ANCOVA, effect size at a visit will be calculated as the absolute value of the LS means difference from Placebo divided by the model estimate of the pooled standard error (the standard error of the LS mean difference divided by the square root of the sum of inverse sample sizes of change from baseline at each treatment group).

16.2. SECONDARY EFFICACY

The secondary efficacy analyses will be performed for the ITT population.

16.2.1. SECONDARY EFFICACY VARIABLE & DERIVATIONS

The secondary efficacy variable is change from baseline in Clinical Global Impression Bipolar Version, Severity of Illness (CGI-BP-S) Score (Depression). The CGI-BP-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 greater illness severity. The CGI-BP-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). CGI-BP-S is assessed at Visit 2/Baseline/Week 0, Visit 3/Week 1, Visit 4/Week 2, Visit 5/Week 4, and Visit 6/Week 6.

16.2.2. MISSING DATA METHODS FOR SECONDARY EFFICACY VARIABLE(S)

The secondary efficacy variable, change from baseline in CGI-BP-S depression score at Week 6, will be set to missing if the CGI-BP-S depression score at Week 6 is missing or baseline value is missing. The same applies to all other visits. The analysis of the secondary efficacy variable will use a MMRM which makes an MAR assumption to the missing secondary efficacy endpoints. Early termination data will be mapped as described in Section 6.3. Missing data will not be imputed.

16.2.3. ANALYSIS OF SECONDARY EFFICACY VARIABLE

The secondary efficacy estimand for this study is the difference in means of change from baseline in the CGI-BP-S depression score at Week 6 comparing all SEP-4199 CR treatment group to placebo in ITT population with major depressive episode associated with bipolar I disease (bipolar I disorder) defined through the inclusion/exclusion criteria, should a subject be able to tolerate and adhere to treatment up to Week 6.

The efficacy of SEP-4199 CR in terms of the CGI-BP-S depression score will be evaluated using the following null hypothesis:

> H_2: There is no difference in mean change from baseline at Week 6 on the CGI-BP-S depression score in the all SEP-4199 CR treatment group compared to Placebo.

The alternative hypothesis for the null hypothesis is that there is a difference.

The secondary efficacy endpoint (change from Baseline in CGI-BP-S depression score at Week 6) will be analyzed with an MMRM model similar to the one used in the primary efficacy analysis (adjusted with the corresponding Baseline). For the secondary efficacy endpoint, two models are used: one for all SEP-4199 CR versus placebo, one for each SEP-4199 CR dose group versus placebo group. The LS mean treatment difference of change from Baseline at Week 6, the 2-sided 95% CIs, and the associated p-value will be calculated based on the model.

The CGI-BP-S mania and overall bipolar illness scores will only be presented in listings.

16.2.4. SUPPORTIVE ANALYSIS OF SECONDARY EFFICACY VARIABLE

ANCOVA analysis as described in Section 16.1.4 will also be produced for the secondary endpoint.

16.3. ADDITIONAL EFFICACY ENDPOINTS ANALYSES

All the additional efficacy endpoints will be analyzed using the ITT population.

16.3.1. ADDITIONAL EFFICACY VARIABLE DERIVATIONS

16.3.1.1. MADRS-6 Core Symptoms Subscale Score

The MADRS-6 core symptoms subscale score is the six-item MADRS subscale score, calculated as the sum of the following items: Apparent sadness, reported sadness, inner tension, lassitude, inability to feel, and pessimistic thoughts (item numbers 1, 2, 3, 7, 8, and 9, and possible score range between 0 to 36). MADRS-6 is assessed at Visit 1/Screening, Visit 2/Baseline/Week 0, Visit 3/Week1, Visit 4/Week 2, Visit 5/Week 4, and Visit 6/Week 6.

16.3.1.2. Hamilton Rating Scale for Anxiety (HAM-A)

The HAM-A is a rating scale developed to quantify the severity of anxiety symptomatology. The 14 items are: anxious mood, tension, fears, insomnia, intellectual, depressed mood, somatic (muscular), somatic (sensory), cardiovascular symptoms, respiratory symptoms, gastrointestinal symptoms, genitourinary symptoms, autonomic symptoms, and behavior at interview. Each of the 14 items is rated on a 5-point scale, ranging from 0 (not present) to 4 (very severe). The HAM-A total score is the sum of the 14 individual items and ranges from 0 to 56. A higher score is associated with a greater degree of anxiety. A score of ≥14 has been suggested to indicate clinically significant anxiety (Rush 2000). If 1 or more items are missing at a visit, the total score will be set to missing. HAM-A is assessed at Visit 2/Baseline/Week 0, Visit 3/Week 1, Visit 4/Week 2, Visit 5/Week 4, and Visit 6/Week 6.

16.3.1.3. Quick Inventory of Depressive Symptomatology – Self-Report 16-Item (QIDS-SR16)

The QIDS-SR16 is a 16-item self-report measure of depressive symptomatology which uses a computerized assessment interface for administration. The scoring system for the QIDS-SR16 converts responses from 16 separate items into nine DSM-IV symptom criterion domains. Each item is rated 0 to 3. For symptom domains that consist of more than one item, the highest score among the items relevant to the given domain is taken. The total score equals the sum of the nine individual domain scores and ranges from 0 to 27. Higher score indicates worse depression symptoms. The nine domains comprise: depressed mood (Item 5); concentration/decision making (Item 10); self outlook (Item 11); suicidal ideation (Item 12); decreased interest (Item 13); decreased energy (Item 14); sleep disturbance (initial, middle, and late insomnia or hypersomnia) (Items 1 to 4); appetite/weight disturbance (Items 6 to 9); and psychomotor disturbance (Items 15 and 16). The total score will be set to missing if one or more domain scores are missing at a visit. The appetite/weight disturbance symptom domain score will be set to missing if three or more items are missing at a visit, given the planned skip pattern of items 6 to 9, and the remaining symptom domain scores will be set to missing if one or more items within the respective domain are missing. QIDS-SR16 is assessed at Visit 2/Baseline/Week 0 and Visit 6/Week 6.

16.3.1.4. Sheehan Disability Scale (SDS)

Three items are self-rated using an 11-point visual analog scale ranging from 0 to 10 to assess disability across three domains: work/school, social life, and family life; the three items will be summarized individually in addition to the SDS total score. The SDS total score is calculated as the sum of the 3 items and ranges from 0 (unimpaired) to 30 (highly impaired). If one or more items are missing at a visit, as can occur when a subject opts out of the work/school item because it does not apply, the

authors of the scale recommend setting the total score to missing (Sheehan 2008). Number of days lost due to symptoms, and number of days underproductive due to symptoms are also collected. SDS is assessed at Visit 2/Baseline/Week 0 and Visit 6/Week 6.

16.3.1.5. EuroQoL 5-D (EQ-5D-5L)

The EQ-5D-5L is a standardized instrument developed by the EuroQol Group as a measure of health-related quality of life that can be used in a wide range of health conditions and treatments. The EQ-5D consists of a descriptive system and the EQ VAS.

The descriptive system comprises five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension has 5 levels: no problems, slight problems, moderate problems, severe problems and extreme problems, corresponding to scores of 1 to 5, respectively. The patient is asked to indicate his/her health state by ticking the box next to the most appropriate statement in each of the five dimensions. This decision results in a 1-digit number that expresses the level selected for that dimension. The digits for the five dimensions can be combined into a 5-digit number that describes the patient's health state. An index value (a weighted scoring of the 5 dimension scores with a possible range from 0 to 1) will be assigned to each observed health state using the US value set as defined in APPENDIX 9.

The EQ VAS records the patient's self-rated health on a vertical visual analogue scale, where the endpoints are labeled 'The best health you can imagine' and 'The worst health you can imagine'. The VAS can be used as a quantitative measure of health outcome that reflect the patient's own judgment. EQ-5D-L is assessed at Visit 2/Baseline/Week 0 and Visit 6/Week 6.

16.3.1.6. Snaith-Hamilton Pleasure Scale (SHAPS)

The SHAPS (Snaith 1995) is a subject-reported questionnaire that will be administered at Visit 2/ Baseline/ Week 0 and Visit 6/Week 6.

The SHAPS is a 14 item scale that assesses 4 domains of hedonic experience: interests/pastimes, social interaction, sensory experience, and food/drink. Subjects are asked to respond based on their ability to experience pleasure in the past few days. Each of the items has a set of four response categories: strongly disagree, disagree, agree, or strongly agree, with either of the Disagree responses receiving a score of 1 and either of the Agree responses receiving a score of 0. Thus, the SHAPS total score is scored as the sum of the 14 items and ranges from 0 to 14. A higher SHAPS total score indicated higher levels of present state of anhedonia.

16.3.1.7. MADRS responders

A MADRS responder is defined as achieving a ≥ 50% reduction from the baseline total score at Week 6 and Week 6 (LOCF). The MADRS total score percentage change will be defined as (value at postbaseline visit – baseline value) x 100 / (baseline value). Subjects having a negative percentage change indicate improvement in MADRS total score.

16.3.1.8. MADRS remission

A MADRS remission is defined as a MADRS total score of ≤ 12 at Week 6 and Week 6 (LOCF).

16.3.2. Missing Data Methods for Other Efficacy Variable(s)

Any individual missing item in any scale will not be imputed.

For MADRS-6, HAM-A, SDS, SHAPS, and YMRS, if 1 or more items are missing at a visit, the total score will be set to missing. Similarly, if an individual domain score for EQ-5D-5L is missing, the index value cannot be calculated.

For QIDS-SR16, the total score will be set to missing if 1 or more domain scores are missing at a visit. The appetite/weight disturbance symptom domain score will be set to missing if 3 or more items are missing at a visit, given the planned skip pattern of items 6 to 9, and the remaining symptom domain scores will be set to missing if 1 or more items within the respective domain are missing.

16.3.3. Analysis of Additional Efficacy Variables

16.3.3.1. Change from baseline in MADRS-6 subscale score at Week 6

The MADRS-6 subscale score and change from baseline will be summarized by time point for each treatment group. Change from baseline at Week 6 will be analyzed by the MMRM method described for primary endpoint with baseline MADRS-6 subscale score as covariate (see Section 16.1.3).

16.3.3.2. Change from baseline in HAM-A total score at Week 6

The HAM-A total score and change from baseline will be summarized by time point for each treatment group. Change from baseline at Week 6 will be analyzed by the MMRM method described for primary endpoint with baseline HAM-A total score as covariate (see Section 16.1.3).

16.3.3.3. Change from baseline in QIDS-SR16 total score at Week 6

The QIDS-SR16 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 the ANCOVA method described for primary endpoint with baseline QIDS-SR16 total score as covariate (see Section 16.1.4).

16.3.3.4. Change from baseline in SDS total score at Week 6

The SDS total score and change from baseline will be summarized by time point for each treatment group. Change from baseline at Week 6 will be analyzed by the ANCOVA model described in Section 16.1.4 with baseline SDS total score as covariate.

16.3.3.5. Change from baseline in EQ-5D-5L Scores at Week 6

The EQ dimension scores, VAS score, and index score and change from baseline will be summarized by time point for each treatment group. Change from baseline of EQ VAS and index score at Week 6 will be analyzed by the ANCOVA model described in Section 16.1.4 with baseline EQ VAS and baseline index score as a covariate, respectively.

16.3.3.6. Change from baseline in SHAPS total score at Week 6

The SHAPS total score and change from baseline will be summarized by time point for each treatment group. Change from baseline at Week 6 will be analyzed by the ANCOVA model described in Section 16.1.5.1 with baseline SHAPS total score as covariate.

16.3.3.7. MADRS responders at Week 6

The MADRS total score responder proportions at Week 6 will be compared between the two treatment groups using logistic regression. This model will include treatment and region as fixed effects and

baseline MADRS score as a covariate. If the model fails to converge, the model will be attempted without region first, and if still not converging then without baseline MADRS total score. Effect size and Number Needed to Treat (NNT) will be provided.

Between group effect size will be defined as: Natural Log(odds ratio)xsqrt(3)/pi

The NNT will be derived as 1/Risk Reduction where Risk Reduction (RR) = (all SEP-4199 CR treatment group – placebo response rate). The NNT results will be provided in whole numbers with any fractional values rounded up to the nearest whole number.

The 95% CI will also be presented (only when the lower and upper limits are of the same sign) and are computed by taking the reciprocal of the 95% lower and upper bound of the RR. The lower confidence limit will be rounded down to the largest integer value that is less than the computed estimate, and the upper confidence limit will be rounded up to the smallest integer value that is greater than the computed estimate.

16.3.3.8. MADRS remission at Week 6

Remission incidence rate, defined as MADRS total score ≤12, will be analyzed similarly to the MADRS responder analysis (see Section 16.3.3.7).

17. SAFETY OUTCOMES

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

There will be no statistical comparisons between the treatment groups for safety data, unless otherwise specified within the relevant section.

17.1. PRETREATMENT AND ADVERSE EVENTS

Pretreatment events and AEs will be coded using the MedDRA central coding dictionary, Version 24.1 or higher.

Adverse events are untoward medical occurrences that occurred on or after the first dose of study medication, with a missing start date and a stop date on or after first dose of study medication, or with both a missing start and stop date. Untoward medical occurrences that started between informed consent and prior to the first dose of study medication are pre-treatment events.

Whenever available, the time information should be accounted for in the derivation of AEs vs. pretreatment 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 drug will be considered AEs; those that started before the day of the first dose of study drug will be considered pretreatment.

See APPENDIX 2 for handling of partial dates for AEs. In the case where it is not possible to define an untoward medical occurrence as an AE or pretreatment event, it will be classified by the worst case, ie, AE.

An overall summary of the incidence of AEs within each of the categories described in the following sections will be provided as specified in the templates, including summary of SAEs. The overall incidence summary will also be provided for AEs related to study drug.

Listings will be provided for all AEs, AEs leading to discontinuation of study drug, AEs leading to discontinuation from the study, serious adverse events (SAEs), and AEs leading to death. A listing for pretreatment events will also be presented.

For incidence summaries, each subject will be counted only once within each SOC and PT. Adverse Events will be sorted alphabetically by SOC and then by decreasing frequency of PT within each SOC based on the All SEP-4199 CR treatment group. For the incidence summary table by SOC and PT, each subject will be counted only once within each SOC and PT, and AEs will be sorted alphabetically by SOC, and then by decreasing frequency of PT within each SOC based on the All SEP-4199 CR group. If not otherwise specified, all summaries will present incidence (number of subjects and percentages) and number of events.

17.1.1. ALL AEs

Adverse events will be summarized by SOC and PT for AE incidence and number of events. A listing of all AEs will be presented.

Adverse events will also be presented by maximum severity and by strongest relationship to the study drug as specified in the sections below.

17.1.1.1. Severity

Severity is classified as mild/ moderate/severe (increasing severity). Adverse events with a missing severity will be summarized as severe. 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.

17.1.1.2. Relationship to Study Drug

Relationship to study drug, as indicated by the investigator, is classed as “not related”/ “possible”/ “probable”/ “definite” (increasing strength of relationship). A “related” AE is defined as an AE with a relationship to the study drug of “possible”, “probable” or “definite”. A “not related” AE is defined as an AE with a relationship to the study drug of “not related”. Adverse events with a missing relationship to the study drug will be regarded as “related” to the study drug. If a subject reports the same AE more than once within the same SOC/PT, the AE with the strongest relationship to study drug will be used in the corresponding relationship summaries. For this summary, AEs will be presented in 2 categories, related and not related, by SOC and PT.

17.1.2. AEs LEADING TO DISCONTINUATION FROM STUDY TREATMENT PERIOD

Adverse events leading to discontinuation from the study treatment period are AEs with “Caused Study Discontinuation” = “Yes” on the AE CRF page for subjects with reason for discontinuation as AE on the Study Disposition CRF page or AE action taken=”Drug Withdrawn”. A summary of AEs leading to discontinuation from the study treatment period by SOC and PT will be presented. A listing of AEs leading to discontinuation from the study treatment period will be presented.

17.1.3. SERIOUS ADVERSE EVENTS AND DEATH

Serious AEs are those AEs recorded as “Serious” on the AE CRF page. AEs leading to death are those AEs which are recorded as having an outcome of “Fatal” on the AE CRF page. Summaries of SAEs (including deaths) by SOC and PT will be prepared. A listing of SAEs (including deaths) will be presented.

17.1.4. ADVERSE EVENTS OF SPECIAL INTEREST (AESI)

A list of preferred terms that are to be combined for assessment of adverse events of special interest (AESI), eg, hyperprolactinemia-related AEs, details of definition are provided in Appendix 5.

AEs potentially associated with the study medication SEP-4199 CR:

- Extrapyramidal symptoms (EPS): SMQ Extrapyramidal symptoms and 4 sub-classes (broad),
- QT prolongation: SMQ Torsade de Pointes/QT prolongation (narrow, broad),

AESI will be summarized by AESI category and preferred term for overall subjects. AE SMQs and clustered PTs will be summarized by category (and sub-category, when applicable) and preferred term.

17.2. LABORATORY EVALUATIONS

All summaries will be based on the SAF population. Laboratory data to be reported for this study include Hematology, Serum Chemistry (including lipid panel and thyroid panel), Urinalysis, Urine drug screening, Hemoglobin (HbA1c), Serum Prolactin, Serology, Serum Insulin, C-Reactive Protein (CRP), and urine pregnancy test. A list of laboratory assessments to be included in the outputs is included in section 21 of the CSP. Urine drug screening, serum FSH, serology, and urine pregnancy test will only be listed.

Presentations will use both international system of units (SI) and conventional (CV) units.

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

The following summaries will be provided for laboratory data:

- By-visit summary of observed values and changes from baseline for continuous data in hematology, chemistry, and urinalysis.
- By visit summary of the number and percentage of subjects in each outcome category for categorical data in urinalysis (if applicable).
- Shift in laboratory results (chemistry, hematology, urinalysis) from baseline to Week 6 according to the reference range criteria. The normal reference ranges from the central laboratory will be used to determine whether the laboratory test value is below, within, or above the normal range. Urinalysis will be presented as normal or abnormal (including low, high, abnormal) in the shift table.
- Number and percentage of subjects with at least one potentially clinically significant (PCS) postbaseline laboratory value (see Section 17.2.1 and APPENDIX 7) will be presented by treatment group. The period of evaluation is the double-blind treatment period, including unscheduled visits. Subjects will be represented in the count of a particular PCS category if they have met that criteria at least once postbaseline, regardless of their baseline value.

Serum prolactin results will be summarized overall and separately by gender. Glucose, insulin, and lipid panel results, along with a derived variable for homeostasis model assessment of insulin resistance (HOMA-IR) will be summarized for fasting only and overall (fasting, non-fasting, or fasting status unknown combined) status. HOMA-IR will be derived based on glucose and insulin results as: HOMA-IR = Glucose (mg/dL) x Insulin (mU/L) / 405 (Matthews 1985). The following conversion factors will be used if needed:

Glucose (mg/dL) = Glucose (mmol/L) × 18.015588;
Insulin (mU/L) = Insulin (pmol/L) × (1/6).

HbA1c will also be summarized.

All laboratory data will be provided in data listings, with the values outside the reference ranges flagged.

17.2.1. Laboratory Reference Ranges and Potentially clinically significant (PCS) Criteria

Quantitative laboratory measurements will be compared with the relevant laboratory reference ranges in SI units and categorized as:

- Low: Below the lower limit of the laboratory reference range.
- Normal: Within the laboratory reference range (upper and lower limit included).
- High: Above the upper limit of the laboratory reference range.

For laboratory parameters with categorical outcomes, measurements will be compared with the relevant laboratory reference values and categorized as normal or abnormal.

In addition to the high and low quantitative laboratory assignments (as identified by means of the laboratory reference ranges), Potentially clinically significant (PCS) quantitative safety (and other) laboratory assessments will also be identified in accordance with the predefined Potentially clinically significant (PCS) criteria as presented in APPENDIX 7.

17.3. ECG Evaluations

Results from the central ECG 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 the SAF population:

- By visit summary of observed values and changes from baseline (for quantitative measurements)
- By visit and overall (including unscheduled visits) summary of ECG overall assessment results. Percentage of subjects will be based on the number of subjects with ECG overall assessment available at the given time point. Any unscheduled ECG that occurs after first dose will be included in the overall post-treatment summaries.
- Shift in ECG overall assessments from baseline to Week 6.
- Number and percentage of subjects with QTc levels in each of the QTc categories

The number and percentage of subjects with QTc values in the following QTc interval prolongation categories will be identified, same criteria apply to both QTcF and QTcB:

- > 450 msec for males \ >470 msec for females at any postbaseline time point (including unscheduled visits) not present at baseline
- > 480 msec for male or females at any postbaseline time point (including unscheduled visits) not present at baseline
- > 500 msec for males or females at any postbaseline time point (including unscheduled visits) not present at baseline
- ≥ 30 msec and < 60 msec increase from baseline for at least one postbaseline measurement (including unscheduled visits)
- ≥ 60 msec increase from baseline for at least one postbaseline measurement (including unscheduled visits)

All ECG parameters, overall interpretation, and findings will be provided in data listings. A listing displaying ECG values for subjects with at least 1 prolonged QTc will be produced.

17.3.1. ECG SPECIFIC DERIVATIONS

The following 3 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 may 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

$$RR\ (\text{msec}) = 1000 * \frac{60}{HR\ (bpm)}$$

17.3.2. ECG Potential Clinically Significant Criteria

Potentially clinically significant ECG measurements will be identified in accordance with the following predefined PCS criteria:

Table 5: Predefined ECG PCS criteria

ECG Parameter	PCS Low	PCS High
Heart Rate (beats/min)	< 50	> 100
PR Interval (msec)	--	> 210
QRS Duration (msec)	--	> 120

17.4. 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)
- Waist circumference (cm)

The following summaries will be provided for vital signs data for the SAF population:

- Observed value and change from baseline by visit including the follow-up visit.
- BMI Categories will be described by visit.
- Number and percentage of subjects with at least one PCS vital sign value (see Section 17.4.1) post Baseline. The period of evaluation is the double-blind treatment period, including unscheduled visits. Subjects will be counted in a particular PCS category if they met that PCS criteria at least once post Baseline, regardless of their Baseline value.
- The number and percentage of subjects with orthostatic hypotension and orthostatic tachycardia will be summarized by treatment for baseline and the overall postbaseline period, as well as by visit. As specified in Section 6.3, any orthostatic hypotension or tachycardia events that occurred at the early termination visit will be assigned to the next planned visit. 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 and diastolic blood pressures measured in the supine position, respectively. Orthostatic tachycardia is defined as a heart rate increase of at least 20 beats per minute (bpm) after the subject was standing for at least 2 to 4 minutes, compared to the heart rate measured in the supine position and a heart rate > 100 bpm after the subject was standing for at least 2 to 4 minutes. Standing vital signs will be used as collected on the CRFs.

All vital signs data will be provided in a data listing. In addition, a separate listing will be generated to present the vital signs data that met the PCS criteria. All occurrences of orthostatic hypotension and orthostatic tachycardia will also be presented in a listing.

17.4.1. Vital Signs Potentially Clinically Significant Criteria

Potentially clinically significant vital sign measurements will be identified in accordance with the following predefined PCS criteria:

Table 6. Predefined Potentially Clinically Significant vital sign criteria.

Parameter Name	Low	Decrease from Baseline	High	Increase from Baseline
Systolic BP	≤ 90 mmHg	≥ 20 mmHg	≥ 180 mmHg	≥ 20 mmHg
Diastolic BP	≤ 50 mmHg	≥ 15 mmHg	≥ 105 mmHg	≥ 15 mmHg
Pulse rate	≤ 50 bpm	≥ 15 bpm	≥ 120 bpm	≥ 15 bpm
Respiration Rate	≤ 10 breaths/min	≥ 50%	≥ 25 breaths/min	≥ 50%
Weight		≥ 7%		≥ 7%
Temperature	N/A		≥ 38.3 °C	≥ 0.8°C increase from

				baseline

Note: A vital sign measurement is PCS Low if it is below the specified lower limit and decreased from baseline orPCS High if it is above the specified upper limit and increased from baseline.

17.5. PHYSICAL AND NEUROLOGICAL EXAMINATION

As all physical and neurological findings will be recoded as medical history or AEs, no specific analysis of physical and neurological examination will be performed.

17.6. OTHER SAFETY ASSESSMENTS

17.6.1. YOUNG MANIA RATING SCALE (YMRS)

The YMRS is an 11-item instrument used to assess the severity of mania in patients with a diagnosis of bipolar disorder. The 11 items are: Elevated Mood, Increased Motor Activity Energy, Sexual Interest, Sleep, Irritability, Speech (Rate and Amount), Language-Thought Disorder, Content, Disruptive-Aggressive Behaviour, Appearance and Insight. The YMRS is a clinician-rated assessment. Ratings are based on patient self-reporting, combined with clinician observation (accorded greater score). Seven items are rated on a 5-point scale, ranging from 0 to 4, and four items are rated on a 9-point scale, ranging from 0 to 8. The YMRS total score is the sum of the 11 individual items and ranges from 0 to 60. A higher score is associated with a greater severity of mania. If one or more items are missing at a visit, the total score will be set to missing. YMRS is assessed at Visit 1/Screening, Visit 2/Baseline/Week 0, Visit 3/Week 1, Visit 4/Week 2, Visit 5/Week 4, Visit 6/Week 6, and Visit 7/Follow-up/Week 7.

The YMRS total score and change from baseline will be summarized by time point for each treatment group. Change from baseline over time will be analyzed by the MMRM method described for primary endpoint with baseline YMRS total score as covariate (see Section 16.1.3).

17.6.2. TREATMENT-EMERGENT MANIA

Treatment-emergent mania is defined as a YMRS score ≥ 16 on any 2 consecutive post-Baseline visits or at the final assessment, or an AE of mania or hypomania during the double blind treatment period. If a subject has no reported AE of mania or hypomania and has no post-Baseline assessment of YMRS, incidence of treatment-emergent mania will be set to missing. The mania indicator will be set to 1 if the subject exhibits post-baseline mania, and 0 if the subject does not experience post-baseline mania and has at least 1 non-missing post-baseline YMRS total score, and missing otherwise.

Number and percentage of patients with treatment-emergent mania will be summarized by treatment group for the SAF population. The proportion of subjects with treatment-emergent mania will be analyzed using the logistic regression model including factors of treatment, region, and Baseline YMRS total score as a covariate. Wald chi-square p-values for tests of significance associated with each effect (treatment, geographic region, and baseline YMRS total score) will also be presented. This analysis will be performed for the SAF population.

17.6.3. 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, past 3 months (90 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 is assessed at Visit 1/Screening, Visit 2/Baseline/Week 0, Visit 3/Week 1, Visit 4/Week 2, Visit 5/Week 4, Visit 6/Week 6, and Visit 7/Follow-up/Week 7.

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 “postbaseline” are defined as follows.

Time point	Study Visit	C-SSRS Version	Derivation Rule
Baseline	Screening/Visit 1	Baseline/Screening – Past 3 Months	Most severe outcome

	Visit 2/Baseline*	Since Last Visit	
Post-baseline	All postbaseline visits up to and including Visit 6/Week 6, including unscheduled visits	Since Last Visit	Most severe outcome

* Note: The Visit 2/Baseline C-SSRS assessment must be administered prior to the first dose of study medication in order to be used in the C-SSRS Baseline derivation.

C-SSRS composite endpoints will be derived for each time point of interest (ie, baseline, postbaseline, and each study visit) 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 Suicidal Ideation or Behavior: 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 (ie, baseline, postbaseline, each study visit including follow-up visit) 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.

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 1 to 5 for frequency and duration, and from 0 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.

The number and percentage of subjects with any suicidality, any suicidal ideation, and any suicidal behavior in the SAF population will be presented for:

- Baseline (as defined above)
- Postbaseline (as defined above)
- Each scheduled study visit: Screening (lifetime; past 3 months), Week 1, Week 2, Week 4, Week 6.

Shift in suicidal ideation score from Baseline to the overall post-Baseline period, to each of the scheduled visits will be presented.

The following comparative endpoints will also be presented, where “Treatment emergence” is used for outcomes that include events that first emerge or worsen, and “Emergence” is used for outcomes that include events that first emerge.

- Treatment-emergent suicidal ideation compared to recent history: An increase in the maximum suicidal ideation score during treatment from the maximum suicidal ideation category at Baseline (C-SSRS scales taken during the specified pre-treatment period; excludes “lifetime” scores from the Baseline C-SSRS scale or Baseline/Screening C-SSRS scale).

- Treatment-emergent serious suicidal ideation compared to recent history: An increase in the maximum suicidal ideation score to 4 or 5 on the C-SSRS during treatment from not having serious suicidal ideation (scores of 0-3) at Baseline (C-SSRS scales taken during the specified pre-treatment period; excludes "lifetime" scores from the Baseline C-SSRS scale or Baseline/Screening C-SSRS scale).
- Emergence of serious suicidal ideation compared to recent history: An increase in the maximum suicidal ideation score to 4 or 5 on the C-SSRS during treatment from no suicidal ideation (scores of 0) at Baseline (CSSRS scales taken during the specified pre-treatment period; excludes "lifetime" scores from the Baseline C-SSRS scale or Baseline/Screening C-SSRS scale).
- Improvement in suicidal ideation at a time point of interest compared to Baseline: An improvement in this endpoint can be considered as a decrease in suicidal ideation score at the time point of interest (e.g., the last measurement during treatment) from the baseline measurement (e.g., the measurement taken just prior to treatment). This analysis should only be performed for studies in which a baseline C-SSRS can be defined (i.e., having improvement from the worse event over a lifetime is not clinically meaningful).
- Treatment-emergent suicidal behavior compared to all prior history: The occurrence of suicidal behavior (categories 6-10) during treatment from not having suicidal behavior (categories 6-10) prior to treatment (includes "lifetime" and/or "screening" scores from the Baseline C-SSRS scale, Screening C-SSRS scale, or Baseline/Screening C-SSRS scale, and any "Since Last Visit" from the Since Last Visit C-SSRS scales taken prior to treatment).

Responses to each question will be listed.

17.6.4. ABNORMAL INVOLUNTARY MOVEMENT SCALE (AIMS)

The AIMS is a clinician-rated assessment of abnormal movements consisting of unobtrusive observation of the subject at rest (with shoes removed) and several questions or instructions directed toward the subject. It contains seven items related to: facial, lip, jaw, and tongue movements, upper and lower extremity movements, and trunk movements. Three other items assess the subject at a global level, and two items assess dental status. The individual items on the AIMS are rated from 0 = 'None' to 4 = 'Severe'. The (non-global) total AIMS score is the sum of items 1 through 7. The possible range for the total AIMS score is 0 to 28. Items 8 through 12 are not included in the total AIMS score. Higher values of the total AIMS score indicate increased severity in abnormal movement. If 1 or more of the AIMS total score items are missing at a visit, the score will be set to missing. Item 8, representing global severity score, will be summarized separately. AIMS is assessed at Visit 2/Baseline/Week 0 and Visit 6/Week 6.

All summaries and analyses will be based on the SAF population. Descriptive statistics for observed value and change from baseline will be displayed at week 6 and Week 6 (LOCF) by treatment group. AIMS total score will be analyzed using ANCOVA based on the change from baseline to Week 6 and Week 6 (LOCF) similar to the primary efficacy endpoint, with the respective baseline values as covariate (see Section 16.1.4).

The AIMS total score at Week 6 and Week 6 (LOCF) will also be classified as 'abnormal' if either: at least two items have a response of 'mild' or higher; or at least one item has a response of 'moderate' or higher. Otherwise, non-missing total scores will be classified as 'normal'. This is a modification of the Schooler-Kane Criteria for Tardive Dyskinesia (Schooler 1982). Shifts from baseline will be summarized at Week 6 and Week 6 (LOCF), postbaseline overall (based on maximum severity), and treatment group.

Frequency distribution of AIMS global severity score will be provided at Week 6 and Week 6 (LOCF) and treatment group. Postbaseline AIMS global severity scores will be classified as 'worsened', 'unchanged', or 'improved', relative to a subject's baseline score. A higher score than that of baseline

would be classified as 'worsened'. Conversely, a lower score would be classified as 'improved'. These postbaseline changes will be summarized at Week 6 and Week 6 (LOCF) and treatment group.

17.6.5. BARNES AKATHISIA RATING SCALE (BARS)

The BARS is a rating scale geared toward assessment of neuroleptic-induced akathisia, though it can be used to measure akathisia associated with other drugs as well. The BARS consists of four items, including one item assessing objective restlessness, two items targeting subjective restlessness (awareness and related distress), and one global clinical assessment item. All items are anchored and utilize a 4-point scale, except for the global rating which has a 6-point scale (from absence of akathisia through severe akathisia). The subjective and objective items are summed to yield a total score. The BARS total score is the sum of items 1 through 3 and ranges from 0 to 9. Higher values of the BARS total score indicate higher severity of akathisia. If one or more of items 1 to 3 at a visit are missing the total will not be calculated. BARS is assessed at Visit 2/Baseline/Week 0 and Visit 6/Week 6.

All summaries and analyses will be based on the SAF population. Descriptive statistics for observed value and change from baseline in BARS total score and global clinical assessment score will be displayed at each visit by treatment group. The BARS total score and global clinical assessment score will be analyzed using ANCOVA based on the change from baseline to Week 6 similar to the primary efficacy endpoint, with the respective baseline values as covariate (see Sections 16.1.4).

Categorical scores for the four BARS items will be summarized with number and percentage of observations at Week 6 and Week 6 (LOCF) and treatment group.

The postbaseline BARS Global Clinical Assessment of Akathisia responses will be classified as 'worsened', 'unchanged', or 'improved', relative to a subject's baseline response. A higher score than that of baseline would be classified as 'worsened'. Conversely, a lower score would be classified as 'improved'. Shifts from baseline will be summarized at Week 6 and Week 6 (LOCF) and treatment group.

17.6.6. MODIFIED SIMPSON-ANGUS SCALE (MSAS)

The Modified SAS is a clinician-rated assessment of neuroleptic-induced Parkinsonism consisting of 10 items. Items are anchor-based, rated on a 5-point scale, and address rigidity, gait (bradykinesia), tremor, glabellar tap, and salivation (Simpson 1970). The MSAS total score is defined as the sum of all 10 items and ranges between 0 and 40. Lower values of the MSAS total score indicate milder symptoms. If one or more items are missing at a visit the SAS total score will be set to missing. MSAS is assessed at Visit 2/Baseline/Week 0 and Visit 6/Week 6.

All summaries and analyses will be based on the SAF population. Descriptive statistics for observed value and change from baseline in MSAS total score will be displayed at each visit by treatment group. The MSAS total score will be analyzed using ANCOVA based on the change from baseline to Week 6 similar to the primary efficacy endpoint, with the baseline MSAS total score value as covariate (see Sections 16.1.4).

The MSAS total score at each visit will also be classified as 'abnormal' if it exceeds 3 (Rush 2000). Otherwise, non-missing total scores will be classified as 'normal.' Shifts from baseline will be summarized at Week 6 and Week 6 (LOCF) and treatment group.

17.6.7. PHYSICIAN WITHDRAWAL CHECKLIST (PWC)

The Physician Withdrawal Checklist (PWC) is used to evaluate symptoms of withdrawal after discontinuation of study drug. The scale includes 20 symptoms and each symptom is assessed on a 4

point scale using the following: 0=Not Present, 1=Mild, 2=Moderate and 3=Severe. The score for each question is summed to compute a total score ranging from 0 to 60. If the response to any question is missing, then the total score will be missing.

This checklist will be assessed at the end of treatment visit (Visit 6/Week 6) and Visit 7/Follow-up visit.

All summaries and analyses will be based on the Safety population. To assess withdrawal effect per PWC symptoms, the PWC total score will be summarized using descriptive statistics by visit and treatment group. The number and percentage of subjects with a “Present” response (mild, moderate, or severe) will be provided for each of the 20 PWC items. In addition, responses of mild, moderate and severe will be summarized separately at each visit.

18. PHARMACOKINETIC ANALYSIS

All PK analysis will be performed using PK population.

Blood sample for plasma concentrations of R-amisulpride, S-amisulpride, total amisulpride, and/or plasma prolactin measurement will be collected at Visit 1/Screening (only for plasma prolactin), Visit 2/Baseline/Week 0, Visit 4/Week 2, Visit 6/Week 6, and Visit 7/Follow-up/Week 7. The PK collection times from Predose on Week 0 to Follow-up and deviations and/or any values excluded from analysis will be presented in data listings. Any plasma concentration or summary statistics below the lower limit of quantification (LLOQ) will be represented by “BLQ” (below the limit of quantification) in tables and listings. Amisulpride plasma concentration will not be presented for the placebo group.

R-amisulpride, S-amisulpride, total amisulpride, and plasma prolactin concentrations at each scheduled sample collection time point (including ET and Follow-up, see Section 6.3 for mapping of the ET visit) will be summarized descriptively (n, mean, median, minimum, maximum, coefficient of variation [CV] and if appropriate, geometric mean and geometric CV [GCV]). In addition, if there is at least 1 concentration < LLOQ within a treatment group at a time point, it will be set to ½ LLOQ for summary statistics calculations. Number and percentage of concentrations that are below the LLOQ will be provided for each visit. All PK summaries will be presented by dose.

18.1. DERIVATION

LLOQ is 0.0500 ng/mL for R-amisulpride and S-amisulpride and LLOQ is 1.56 ng/mL for plasma prolactin.

Below derivations apply after concentrations lower then LLOQ are set to ½ LLOQ.

- Coefficient of variation

  100*Standard Deviation/Mean

- Geometric Mean

  Exponential (mean of $\log_e$ transformed data)

- Geometric CV

  Square Root (Exponential (Variance ($\log_e$ transformed data)-1))*100

19. DATA NOT SUMMARIZED OR PRESENTED

Disposition, demographics, and pretreatment events will be listed for screened subjects including the screen failures. A randomized subject's any premedication data from his/her screen failed period will be included in ADaM and in Tables/Listings. Any pre-treatment medical occurrences from the subject's screen failed period will be excluded from ADAE and all other screen failed period data will be excluded from any baseline calculations for the later randomized subject. The information on scales collected on paper is not included in ADaM and in Tables/Listings. Other data that are collected on screen failures will not be presented, but will be available in the clinical study database and SDTM domains.

20. 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.

21. REFERENCES

Buckland ST, Burnham KP, Augustin NH. Model selection an integral part of inference. Biometrics. 1997;53:603–618.

Dmitrienko A, Koch GG. Analysis of clinical trials using SAS: A practical Guide. Second Edition. SAS Press: Cary, NC. 2017.

Dmitrienko A, Molenberghs G, Chuang-Stein C, Offen W. Analysis of Clinical Trials Using SAS: A Practical Guide. SAS Press: Cary, NC. 2005.

Diggle PJ, Heagerty P, Liang KY, Zeger SL. The analysis of Longitudinal Data. Oxford, England: 2nd Edition. Oxford University Press. 2002.

Loebel A, Cucchiaro J, Silva R, Kroger H, Hsu J, Sachs G. Lurasidone monotherapy in the treatment of bipolar I depression: a randomized, double-blind, placebo-controlled study. Am J Psychiatry. 2014 Feb;171(2):160 8.

Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia. 1985;28 (7):412–9.

O'Kelly M, Ratitch B. Clinical Trials with Missing Data: A Guide for Practitioners. John Wiley & Sons. 2014.

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical. Computing, Vienna, Austria. URL https://www.R-project.org/. 2017.

Rush A, Pincus H, First M, Blacker D, Endicott J, Keith S, Phillips K, Ryan N, Smith G, Tsuang M, Widiger T, Zarin D. Handbook of Psychiatric Measures, Washington D.C: American Psychiatric Association. 2000.

Rush J, Pincus H, First M, Blacker D, Endicott J, Keith S, Phillips K, Ryan N, Smith G, Tsuang M, Widiger T, Zarin D. Handbook of Psychiatric Measures. American Psychiatric Association. 2000;554-556.

Schooler NR, Kane JM. Research Diagnoses for Tardive Dyskinesia. Arch. Gen. Psychiatry. 1982;39:486-487.

Sheehan KH, Sheehan DV, Assessing treatment effects in clinical trials with the Discan metric of the Sheehan Disability Scale. International Clinical Psychopharmacology. 2008;23:70-83.

APPENDIX 1. PROGRAMMING CONVENTIONS FOR OUTPUTS

OUTPUT CONVENTIONS

Where applicable, the Appendix_compilation_working_update_2022Final.pdf document – provided by SMPA – will be followed.

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

- The left-hand column should start in column 1. No 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” (eg, 99.99% is presented as “>99.9”); and cases where percent < 0.1% will be presented as “<0.1” (eg, 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 four decimal places, except values <1.0000 but >0.9999 will be presented as '>0.9999' ; and values <0.0001 will be presented as '<0.0001'. Rounding will be applied after the <0.0001 and >0.9999 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 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
Placebo	Placebo

Treatment Group	For Tables, Graphs and Listings
SEP-4199 CR 200 mg	SEP-4199 CR 200 mg (applies to disposition, demographic, some key efficacy and safety tables, and listings)
SEP-4199 CR 400 mg	SEP-4199 CR 400 mg (applies to disposition, demographic, some key efficacy and safety tables, and listings)
All SEP-4199 CR	All SEP-4199 CR (only applies to tables and graphs)

LISTINGS

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

- Randomized/Actual treatment received as applicable, displaying SEP-4199 CR first and then placebo.
- Subject ID,
- Original date/time (where applicable) - listings of AEs, concomitant medications, medical histories etc should be sorted in chronological order, with earliest AE, 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, study medication start date means the date of first dose of study drug.

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 < Study medication start date, then pre-treatment events If start date ≥ Study medication start date, then AE
Known	Partial	If start date < Study medication start date, then pre-treatment events If start date ≥ Study medication start date, then AE
Known	Missing	If start date < Study medication start date, then pre-treatment events If start date ≥ Study medication start date, then AE
Partial, but known components show that it cannot be on or after Study medication start date	Known	Pre-treatment events
Partial, but known components show that it cannot be on or after Study medication start date	Partial	Pre-treatment events
Partial, but known components show that it cannot be on or after Study medication start date	Missing	Pre-treatment events
Partial	Partial	Impute start date as earliest possible date (ie, first day of month if day unknown or 1st January if day and month are unknown). Then: If start date < Study medication start date, then pre-treatment events If start date ≥ Study medication start date, then AE
Partial, could be on or after Study medication start date	Known	If stop date < Study medication start date, then pre-treatment events If stop date ≥ Study medication start date, then AE

START DATE	STOP DATE	ACTION
Partial, could be on or after Study medication start date	Partial	Impute stop date as latest possible date (ie, last day of month if day unknown or 31st December if day and month are unknown). Then: If stop date < Study medication start date, then pre-treatment events If stop date ≥ Study medication start date, then AE
Partial	Missing	Assumed SEP-4199 AE
Missing	Known	If stop date < Study medication start date, then pre-treatment events If stop date ≥ Study medication start date, then AE
Missing	Partial	Impute stop date as latest possible date (ie, last day of month if day unknown or 31st December if day and month are unknown). Then: If stop date < Study medication start date, then pre-treatment events If stop date ≥ Study medication start date, then AE
Missing	Missing	Assumed SEP-4199 AE

ALGORITHM FOR PRIOR / CONCOMITANT / POST-TREATMENT MEDICATIONS:

The concept of "date" below should also include time information whenever available. Time will not be imputed.

START DATE	STOP DATE	ACTION
Known	Known or Ongoing	If stop date < study medication start date, assign as prior. If study medication start date ≤ stop date ≤ end of treatment (date of last dose of study medication), and 1. start date < study medication start date, assign as prior and concomitant. 2. study medication start date ≤ start date ≤ end of treatment assign as concomitant. If stop date > end of treatment or ongoing and 1. start date < study medication start date, assign as prior, concomitant, and post-treatment. 2. study medication start date ≤ start date ≤ end of treatment assigns as concomitant and post-treatment. 3. start date > end of treatment , assign as post-treatment.
Known	Partial	Impute stop date as: • If only the day is unknown, and the month of the stop date is not the same as the month of the the last study visit, then impute as the last day of the month. • If only the day is unknown, and the month of the stop date is the same as the month of the last study visit, then impute as the last study visit date. • If the month and day are unknown, and the year of the stop date is not the same as the year of the the last study visit, then impute as December 31. • If the month and day are unknown, and the year of the stop date is the same as the year of the last study visit, then impute as the last study visit date. Then apply the rule from one of both dates that are known.
Known	Missing(If missing and not ongoing, send a query to DM)	If start date < study medication start date, assign as prior and concomitant, post-treatment. If the study medication start date ≤ start date ≤ end of treatment, assign as concomitant and post-treatment. If start date > end of treatment , assign as post-treatment.
Partial	Known or Ongoing	Impute start date as: *CRF questions: 'Started prior to First Dose?' = Yes; 'Started after last dose of study medication?' = No.* • If only the day is unknown, impute as 1st day of the month. • If the month and day are unknown, impute as January 1. *CRF questions: 'Started prior to First Dose?' = No; 'Started after last dose of study medication?' = Yes.* • If only the day is unknown, impute as the later of (1st day of the month; end of treatment + 1). • If month and day are unknown, impute as the later of (January 1; end of treatment + 1). *CRF questions: 'Started prior to First Dose?' = No; 'Started after last dose of study medication?' = No.* • If only the day is unknown, and the month of start date is the same as the month of the first dose of the study medication then impute as the first dose date of the study medication. • If only the day is unknown, and the month of start date is after the month of the first dose of the study medication then impute as the 1st day of the month. • If the month and day are unknown, and the year of start date is the same as the year of the first dose of the study medication then impute as the first dose date of the study medication. • If the month and day are unknown, and the year of start date is after the year

		of the first dose of the study medication then impute as January 1. Then apply the rule from one of both dates that are known.
Partial	Partial	Impute start date: *CRF questions: 'Started prior to First Dose?' = Yes; 'Started after last dose of study medication?' = No.* • If only the day is unknown, impute as 1st day of the month. • If the month and day are unknown, impute as January 1. *CRF questions: 'Started prior to First Dose?' = No; 'Started after last dose of study medication?' = Yes.* • If only the day is unknown, impute as the later of (first day of the month, end of treatment + 1). • If month and day are unknown, impute as the later of (January 1, end of treatment + 1). *CRF questions: 'Started prior to First Dose?' = No; 'Started after last dose of study medication?' = No.* • If only the day is unknown, and the month of start date is the same as the month of the first dose of the study medication then impute as the first dose date of the study medication. • If only the day is unknown, and the month of start date is after the month of the first dose of the study medication then impute as the 1st day of the month. • If the month and day are unknown, and the year of start date is the same as the year of the first dose of the study medication then impute as the first dose date of the study medication. • If the month and day are unknown, and the year of start date is after the year of the first dose of the study medication then impute as January 1. Impute stop date as: • If only the day is unknown, and the month of the stop date is not the same as the month of the last study visit, then impute as the last day of the month. • If only the day is unknown, and the month of the stop date is the same as the month of the last study visit, then impute as the last study visit date. • If the month and day are unknown, and the year of the stop date is not the same as the year of the last study visit, then impute as December 31. • If the month and day are unknown, and the year of the stop date is the same as the year of the last study visit, then impute as the last study visit date. Then apply the rule from one of both dates are known.
Partial	Missing	Impute start date as: *CRF questions: 'Started prior to First Dose?' = Yes; 'Started after last dose of study medication?' = No.* • If only the day is unknown, impute as 1st day of the month. • If the month and day are unknown, impute 1st January. *CRF questions: 'Started prior to First Dose?' = No; 'Started after last dose of study medication?' = Yes.* • If only the day is unknown, impute as the later of (first day of the month; end of treatment + 1). • If month and day are unknown, impute as the later of (January 1; end of treatment + 1). *CRF questions: 'Started prior to First Dose?' = No; 'Started after last dose of study medication?' = No.* • If only the day is unknown, and the month of start date is the same as the month of the first dose of the study medication then impute as the first dose date of the study medication. • If only the day is unknown, and the month of start date is after the month of the first dose of the study medication then impute as the 1st day of the month. • If the month and day are unknown, and the year of start date is the same as the year of the first dose of the study medication then impute as the first dose date of the study medication. • If the month and day are unknown, and the year of start date is after the year of the first dose of the study medication then impute as January 1.

		Then apply the rule from the one of start dates known and the stop date missing.
Missing(Query to DM and try to confirm the date.)	Known	*CRF questions: 'Started prior to First Dose?' = Yes; 'Started after last dose of study medication?' = No.* • If stop date < study medication start date, assign as prior. • If study medication start date ≤ stop date ≤ end of treatment, then assign prior and concomitant. • If stop date > end of treatment, then assign prior, concomitant, and post-treatment. *CRF questions: 'Started prior to First Dose?' = No; 'Started after last dose of study medication?' = Yes.* • Assign as post-treatment *CRF questions: 'Started prior to First Dose?' = No; 'Started after last dose of study medication?' = No.* • If stop date ≤ end of treatment, then assign concomitant. • If stop date > end of treatment, then assign concomitant, and post-treatment.
Missing(Query to DM and try to confirm the date.)	Partial	Impute stop date as: • If only the day is unknown, and the month of the stop date is not the same as the month of the last study visit, then impute as the last day of the month. • If only the day is unknown, and the month of the stop date is the same as the month of the last study visit, then impute as the last study visit date. • If the month and day are unknown, and the year of the stop date is not the same as the year of the last study visit, then impute as December 31 . • If the month and day are unknown, and the year of the stop date is the same as the year of the last study visit, then impute as the last study visit date. After the End date is imputed follow the Missing/Known category.
Missing(Query to DM and try to confirm the date.)	Missing(Query to DM and try to confirm the date.)	Assign as prior, concomitant, and post-treatment.

PARTIAL DATE IMPUTATION RULES FOR PSYCHIATRICS HISTORY:

Partial dates of first symptoms, initial diagnosis of Bipolar I disorder, onset of last manic episode, and onset of current major depressive episode, will be imputed as follows:
(1) If year and month are known, and if the year or month is before Screening, use the 15th of the month; otherwise, if the month is the month of Screening, use the 1st of the month.
(2) If only year is known, and it is before the year of Screening, use June 30th of that year; If only year is known and it is same as the year of Screening, use Jan 1.
This imputation rule will be applied to calculation of duration of initial symptoms of bipolar I disorder, initial diagnosis of bipolar I disorder, onset of last manic episode, and onset of current major depressive episode.

APPENDIX 3. STATISTICAL MODEL SPECIFICATIONS

MIXED MODEL FOR REPEATED MEASURES (MMRM)

```
Ods graphics on;
PROC MIXED DATA=XXX plots=(studentpanel(marginal)  pearsonpanel(marginal));
        CLASS SUBJID REGION TRTPN WEEK;
        MODEL TOTCFB = TOTBL REGION TRTPN WEEK TRTPN*WEEK / DDFM=KR RESIDUAL OUTPM=OUTRESIDUAL;
        REPEATED WEEK/TYPE=UN SUBJECT=SUBJID;
        LSMEANS TRTPN*WEEK /PDIFF CL;
RUN;
*Where WEEK = Week 1, Week 2, Week 4, Week 6;
Ods graphics off;
```

In case the MMRM model assuming the unstructured covariance variance structure fails to converge, Heterogeneous Toeplitz and Toeplitz, and Compound Symmetry covariance structure will be assumed sequentially. The first covariance structure to yield convergence will be used in the analysis.

```
PROC MIXED empirical;
        CLASS SUBJID REGION TRTPN WEEK;
        MODEL TOTCFB = TOTBL REGION TRTPN WEEK TRTPN*WEEK / RESIDUAL OUTP=OUTRESIDUAL;
        REPEATED WEEK/TYPE=TOEPH  SUBJECT=SUBJID;
        REPEATED WEEK/TYPE=TOEP SUBJECT=SUBJID;
        REPEATED WEEK/TYPE=CS SUBJECT=SUBJID;
        LSMEANS TRTPN*WEEK /PDIFF CL COV;
RUN;
```

Within-group effect size will be calculated as the observed value of LS mean of change from baseline divided by SD, obtained as the SE of the LS Mean multiplied by the square root of the treatment group change from baseline sample size at each visit. Between-group effect size at a visit will be calculated as the observed value of the LS means difference from Placebo divided by the model estimate of the pooled SD at the visit, which is obtained from the square root of the diagonal element, associated at the visit, from the covariance matrix (R matrix of subjects with MADRS total score at all visits). Although the variance-covariance estimates are calculated differently depending on the covariance structure being applied for the model, the diagonal of the R matrix will contain the variance estimate at each visit that can be used for the effect size calculations regardless of the covariance structure.

ANALYSIS OF COVARIANCE (ANCOVA)

```
PROC MIXED;
        BY WEEK;
        CLASS TRTPN REGION;
        MODEL TOTCFB = TRTPN REGION TOTBL;
        LSMEANS TRTPN / PDIFF CL;
RUN;
```

LOGISTIC REGRESSION

```
* Placebo is reference;
PROC LOGISTIC DESCENDING;
        CLASS TRTPN REGION / PARAM=REF;
        MODEL RESP = TRTPN REGION TOTBL;
RUN;
```

APPENDIX 4. INTERNATIONALLY AGREED ORDER OF SYSTEM ORGAN CLASSES

Internationally Agreed Order
Infections and infestations
Neoplasms benign, malignant and unspecified (incl cysts and polyps)
Blood and lymphatic system disorders
Immune system disorders
Endocrine disorders
Metabolism and nutrition disorders
Psychiatric disorders
Nervous system disorders
Eye disorders
Ear and labyrinth disorders
Cardiac disorders
Vascular disorders
Respiratory, thoracic and mediastinal disorders
Gastrointestinal disorders
Hepatobiliary disorders
Skin and subcutaneous tissue disorders
Musculoskeletal and connective tissue disorders
Renal and urinary disorders
Pregnancy, puerperium and perinatal conditions
Reproductive system and breast disorders
Congenital, familial and genetic disorders
General disorders and administration site conditions
Investigations
Injury, poisoning and procedural complications
Surgical and medical procedures
Social circumstances
Product issues

APPENDIX 5. ADVERSE EVENT OF SPECIAL INTEREST (AESI)

During the trial ongoing period and prior to database lock, adverse event customized query defined in this section is used more for AESI reconciliation purposes. Electronic data capture (EDC) system of clinical database contains an AESI tick box, which allow any event term listed below be ticked, thus would encompass any such event in EDC. In that way AESI per customed query (such as hyperprolactinemia-related adverse event) can reconciled for any terms that potentially meet the protocol AESI criteria which have not been identified as such via the AESI tick box (and vice versa). Final definition of each customed query will be finalized right before the DBL and displayed in this SAP before the sign-off. AE preferred terms listed in this section are per MEDDRA version 24.1.

A5.1 Hyperprolactinemia-related AEs (Customized)

"Hyperprolactinemia-related AEs" (customized) is defined as AEs with any of the following preferred terms and all preferred terms that include the word "FRACTURE".

Hyperprolactinemia		
AMENORRHOEA	AMENORRHOEA-GALACTORRHOEA SYNDROME	ANORGASMIA
ANOVULATORY CYCLE	BLOOD OESTROGEN ABNORMAL	BLOOD OESTROGEN DECREASED
BLOOD PROLACTIN ABNORMAL	BLOOD PROLACTIN INCREASED	BONE DEMINERALISATION
BONE DENSITY ABNORMAL	BONE FORMATION TEST ABNORMAL	BONE METABOLISM BIOCHEMICAL MARKER INCREASED
BONE METABOLISM DISORDER	BONE RESORPTION TEST ABNORMAL	BREAST DISCHARGE
BREAST DISCOMFORT	BREAST DISORDER	BREAST DISORDER FEMALE
BREAST DISORDER MALE	BREAST ENGORGEMENT	BREAST ENLARGEMENT
BREAST FIBROSIS	BREAST HYPERPLASIA	BREAST INFLAMMATION
BREAST OEDEMA	BREAST PAIN	BREAST SWELLING
BREAST TENDERNESS	DISTURBANCE IN SEXUAL AROUSAL	EJACULATION DISORDER
ERECTILE DYSFUNCTION	FEMALE ORGASMIC DISORDER	FEMALE SEXUAL AROUSAL DISORDER
FEMALE SEXUAL DYSFUNCTION	FIBROCYSTIC BREAST DISEASE	GALACTOCELE
GALACTORRHOEA	GALACTOSTASIS	GYNAECOMASTIA
HIRSUTISM	HYPERPROLACTINAEMIA	HYPOGONADISM
HYPOMENORRHOEA	INFERTILITY	LIBIDO DECREASED
LIBIDO DISORDER	MACROPROLACTINAEMIA	MALE ORGASMIC DISORDER
MALE SEXUAL DYSFUNCTION	MENSTRUAL DISORDER	MENSTRUATION DELAYED
MENSTRUATION IRREGULAR	METRORRHAGIA	NIPPLE OEDEMA
NIPPLE PAIN	NIPPLE SWELLING	OESTROGEN DEFICIENCY
OESTROGENS TOTAL URINE DECREASED	OLIGOMENORRHOEA	OSTEOPENIA
OSTEOPOROSIS	PAINFUL EJACULATION	PITUITARY AMENORRHOEA
PREMATURE EJACULATION	PROLACTIN-PRODUCING PITUITARY TUMOUR	PSEUDOGYNAECOMASTIA
SEXUAL DYSFUNCTION	SEXUAL INHIBITION	
BREAST ENLARGEMENT FEMALE	HYPOGONADISM FEMALE	HYPOGONADISM MALE
PRIMARY HYPOGONADISM	SECONDARY HYPOGONADISM	INFERTILITY FEMALE
INFERTILITY MALE	LOSS OF LIBIDO	ORGANIC ERECTILE DYSFUNCTION
PSYCHOGENIC ERECTILE DYSFUNCTION		

A5.2 Extrapyramidal Syndrome (EPS) related AEs (Broad)

A "broad" SMQ search includes both the "narrow" scope terms and the additional "broad" scope terms. SMQ "Extrapyramidal syndrome [20000095] (broad)" is defined as AEs with any of the following PTs, which includes all 4 sub-SMQs (Akathisia [20000096], Dyskinesia (including tardive dyskinesia) [20000097], Dystonia [20000098], and Parkinson-like events [20000099])

Akathisia [20000096]		
Narrow		
AKATHISIA		
Broad		
EXTRAPYRAMIDAL DISORDER	HYPERKINESIA	HYPERKINESIA NEONATAL
MOTOR DYSFUNCTION	MOVEMENT DISORDER	PSYCHOMOTOR HYPERACTIVITY
RESTLESSNESS		
Dyskinesia [20000097]		
Narrow		
ATHETOSIS	BALLISMUS	BUCCOGLOSSAL SYNDROME
CHOREA	CHOREOATHETOSIS	DOPAMINE DYSREGULATION SYNDROME
DYSKINESIA	DYSKINESIA NEONATAL	DYSKINESIA OESOPHAGEAL
GRIMACING	OCULOGYRIC CRISIS	PHARYNGEAL DYSKINESIA
PROTRUSION TONGUE	RABBIT SYNDROME	RESPIRATORY DYSKINESIA
TARDIVE DYSKINESIA		
Broad		
ABNORMAL INVOLUNTARY MOVEMENT SCALE	CHRONIC TIC DISORDER	COMPLEX TIC
DROOLING	EXTRAPYRAMIDAL DISORDER	MOTOR DYSFUNCTION
MOVEMENT DISORDER	MUSCLE TWITCHING	PROVISIONAL TIC DISORDER
SECONDARY TIC	TIC	
Dystonia [20000098]		
Narrow		
DYSTONIA	DYSTONIC TREMOR	EMPROSTHOTONUS
MEIGE'S SYNDROME	OCULOGYRIC CRISIS	OPISTHOTONUS
OROMANDIBULAR DYSTONIA	PHARYNGEAL DYSTONIA	PLEUROTHOTONUS
SPASMODIC DYSPHONIA	TORTICOLLIS	TRISMUS
WRITER'S CRAMP		
Broad		
BLEPHAROSPASM	CHRONIC TIC DISORDER	COMPLEX TIC
DROOLING	EXTRAPYRAMIDAL DISORDER	FACIAL SPASM
GAIT INABILITY	LARYNGOSPASM	MOTOR DYSFUNCTION
MOVEMENT DISORDER	MUSCLE CONTRACTIONS INVOLUNTARY	MUSCLE SPASMS
MUSCLE SPASTICITY	MUSCLE TIGHTNESS	MUSCLE TONE DISORDER
MUSCLE TWITCHING	MUSCULOSKELETAL STIFFNESS	OESOPHAGEAL SPASM
OROPHARYNGEAL SPASM	POSTURE ABNORMAL	POSTURING
PROVISIONAL TIC DISORDER	RISUS SARDONICUS	SECONDARY TIC
TIC	TONGUE SPASM	TORTICOLLIS PSYCHOGENIC
UVULAR SPASM		
Parkinson-like events [20000099]		
Narrow		
AKINESIA	BRADYKINESIA	COGWHEEL RIGIDITY
FREEZING PHENOMENON	HYPERTONIA	HYPERTONIA NEONATAL
HYPOKINETIC DYSARTHRIA	MUSCLE RIGIDITY	ON AND OFF PHENOMENON
PARKINSONIAN CRISIS	PARKINSONIAN GAIT	PARKINSONIAN REST TREMOR

PARKINSONISM	PARKINSONISM HYPERPYREXIA SYNDROME	PARKINSON'S DISEASE
PARKINSON'S DISEASE PSYCHOSIS	PROPULSIVE GAIT	RESTING TREMOR
Broad		
ACTION TREMOR	BRADYPHRENIA	DROOLING
DYSPHONIA	EXTRAPYRAMIDAL DISORDER	FINE MOTOR SKILL DYSFUNCTION
GAIT DISTURBANCE	HYPOKINESIA	HYPOKINESIA NEONATAL
LARYNGEAL TREMOR	MICROGRAPHIA	MOBILITY DECREASED
MOTOR DYSFUNCTION	MOVEMENT DISORDER	MUSCLE TONE DISORDER
MUSCULOSKELETAL STIFFNESS	POSTURAL REFLEX IMPAIRMENT	POSTURAL TREMOR
REDUCED FACIAL EXPRESSION	TREMOR	TREMOR NEONATAL
WALKING DISABILITY		

A5.3 Torsade de pointes/QT Prolongation (Broad and Narrow) related AEs

SMQ "Torsade de Pointes/QT prolongation [20000001] (broad and narrow)" is defined as AEs with any of the following PTs:

Torsade de Pointes/QT prolongation [20000001]		
Narrow		
ELECTROCARDIOGRAM QT INTERVAL ABNORMAL	ELECTROCARDIOGRAM QT PROLONGED	LONG QT SYNDROME
LONG QT SYNDROME CONGENITAL	TORSADE DE POINTES	VENTRICULAR TACHYCARDIA
Broad		
ARRHYTHMIC STORM	CARDIAC ARREST	CARDIAC DEATH
CARDIAC FIBRILLATION	CARDIO-RESPIRATORY ARREST	ELECTROCARDIOGRAM REPOLARISATION ABNORMALITY
ELECTROCARDIOGRAM U WAVE INVERSION	ELECTROCARDIOGRAM U WAVE PRESENT	ELECTROCARDIOGRAM U-WAVE ABNORMALITY
LOSS OF CONSCIOUSNESS	SUDDEN CARDIAC DEATH	SUDDEN DEATH
SYNCOPE	VENTRICULAR ARRHYTHMIA	VENTRICULAR FIBRILLATION
VENTRICULAR FLUTTER	VENTRICULAR TACHYARRHYTHMIA	

APPENDIX 6. PREDEFINED POTENTIALLY CLINICALLY SIGNIFICANT (PCS) CRITERIA FOR LABORATORY VALUES

Parameter	PCS Range (SI units)	PCS Range (CV units)
Serium Chemistry		
Sodium	< 130 mmol/L	< 130 mEq/L
	> 150 mmol/L	> 150 mEq/L
Potassium	< 3 mmol/L	< 3 mEq/L
	> 5.5 mmol/L	> 5.5 mEq/L
Chloride	≤ 90 mmol/L	≤ 90 mEq/L
	≥ 118 mmol/L	≥ 118 mEq/L
Calcium	< 1.75 mmol/L	<7.01 mg/dL
	≥ 3.1 mmol/L	≥12.42 mg/dL
Phosphate	< 0.65 mmol/L	<2.01 mg/dL
	> 1.65 mmol/L	>5.11 mg/dL
Bicarbonate	< 15.1 mmol/L	< 15.1 mEq/L
	> 34.9 mmol/L	> 34.9 mEq/L
Magnesium	< 0.4 mmol/L	<0.97 mg/dL
	> 1.23 mmol/L	>2.99 mg/dL
AST	≥ 3 x ULN	≥ 3 x ULN
ALT	≥ 3 x ULN	≥ 3 x ULN
Alkaline Phosphatase	≥ 1.5 x ULN	≥ 1.5 x ULN
GGT (Gamma-Glutamyl Transferase)	≥ 2.5 x ULN	≥ 2.5 x ULN
LDH	≥ 3 x ULN	≥3 x ULN
Creatine Kinase	≥ 3 x ULN	≥ 3 x ULN
Creatinine	≥ 177 umol/L	≥ 2.0 mg/dL
Creatinine Clearance	< 0.48343 mL/s	<28.95 mL/min
BUN	≥ 10.7 mmol/L	≥29.96 mg/dL
Total Bilirubin	>= 34.2 umol/L OR > 2 x ULN	>= 2.0 mg/dL OR > 2 x ULN
Total protein	≤ 45 g/L	≤ 4.5 g/dL
	≥ 100 g/L	≥ 10 g/dL
Albumin	≤ 25 g/L	≤ 2.5 g/dL
Total Cholesterol (fasting)	≥ 7.76 mmol/L	≥ 300 mg/dL
HDL-Cholesterol (fasting)	< 0.78 mmol/L	<30 mg/dL
LDL Cholesterol (fasting)	> 4.14 mmol/L	> 160 mg/dL
Triglycerides (fasting)	> 3.42 mmol/L	> 302.92 mg/dL

Parameter	PCS Range (SI units)	PCS Range (CV units)
Uric acid Male Female	 > 595 umol/L > 476 umol/L	 > 10 mg/dL > 8 mg/dL
Glucose (fasting)	< 2.78 mmol/L	< 50.09 mg/dL
	> 13.9 mmol/L	> 150.45 mg/dL
HbA1c	≥ 0.075	≥ 7.5%
Prolactin	≥ 5 x ULN	≥ 5 x ULN
Hematology		
WBC	≤ 2.8 x10 ^{9}/L	≤ 2.8 x10 ^{3}/µL
	≥ 16 x10 ^{9}/L	≥ 16 x10 ^{3}/µL
Neutrophils (abs)	< 0.5 x 10^{9}/L	< 0.5 x 10^{3}/µL
	> 13.5 x 10^{9}/L	> 13.5 x 10^{3}/µL
Lymphocytes (abs)	> 12 x 10^{9}/L	> 12 x 10^{3}/µL
Monocytes (abs)	> 2.5 x 10^{9}/L	> 2.5 x 10^{3}/µL
Eosinophils (abs)	> 1.6 x 10^{9}/L	> 1.6 x 10^{3}/µL
Basophils (abs)	> 1.6 x 10^{9}/L	> 1.6 x 10^{3}/µL
Neutrophils (relative)	≤ 0.15	≤ 15%
	> 0.85	> 85%
Lymphocytes (relative)	≥ 0.75	≥ 75%
Monocytes (relative)	≥ 0.15	≥ 15%
Eosinophils (relative)	≥ 0.10	≥ 10%
Basophils (relative)	≥ 0.10	≥ 10%
Hemoglobin	Male: ≤ 115 g/L	Male: ≤ 11.5 g/dL
	Female: ≤95 g/L	Female: ≤ 9.5 g/dL
	Male: ≥ 190 g/L	Male: ≥ 19.0 g/dL
	Female: ≥175 g/L	Female: ≥ 17.5 g/dL
Hematocrit	Male: ≤ 0.37	Male: ≤ 37%
	Female: ≤ 0.32	Female: ≤ 32%
	Male: ≥ 0.60	Male: ≥ 60%
	Female: ≥ 0.54	Female: ≥ 54%
RBC	≤ 3.5 x 10^{12}/L	≤ 3.5 x 10^{6}/µL
	≥ 6.4 x 10^{12}/L	≥ 6.4 x 10^{6}/µL
Platelet Count	≤ 75 x 10^{9}/L	≤ 75 x 10^{3}/µL
	≥ 700 x 10^{9}/L	≥ 700 x 10^{3}/µL
Coagulation		
aPTT (sec)	> 1.5 x ULN	> 1.5 x ULN
INR (ratio)	> 1.5 x ULN	> 1.5 x ULN
Urinalysis		
RBC	>25 hpf	>25 hpf
WBC	>25 hpf	>25 hpf

APPENDIX 7. BLINDED DATA REVIEW (BDR) PROCESS

The blinded review of the SEP380-301 data will be conducted prior to unblinding in order to permit the checking and assessment of data for the purposes of identifying additional information for analyses specified in the SAP. This process will be completed prior to unblinding.

This review will be performed by appropriate personnel including, but not limited to a biostatistician and an MD. Relevant outputs will be generated shortly before/after database soft lock in order to facilitate the blinded review process.

Any changes to the planned analysis based on a blinded review of data will be documented with detailed explanation for deviations from the pre-defined plans. Decisions made at this meeting will be documented prior to unblinding to be incorporated into the analysis.

The BDR process will include review of analysis populations. In addition, the IPD review process detailed in the IPD specification document will help identify any IPDs.

APPENDIX 8. **EQ-5D-5L INDEX VALUE**

EQ-5D-5L health states, defined by the EQ-5D-5L descriptive system, will be converted into a single index value using an appropriate value set. For international trials, EuroQol recommends applying a single value set or crosswalk to all study sites. US value set will be used for US, Europe, and Japan for this study. Please refer the following EuroQol website for EQ-5D-5L value sets and further information.

https://euroqol.org/publications/key-euroqol-references/value-sets/

APPENDIX 9.**TABLE OF CONTENTS**

(ITT Population)

Table 14.2.3.6.1 EQ-5D-5L VAS Score – ANCOVA of Change from Baseline at Week 6 and Week 6 (LOCF) (ITT Population)

Table 14.2.3.6.2 EQ-5D-5L Index Value – ANCOVA of Change from Baseline at Week 6 and Week 6 (LOCF) (ITT Population)

Table 14.2.3.6.3 EQ-5D-5L Dimension Score – Observed Value and Change from Baseline at Week 6 and Week 6 (LOCF) (ITT Population)

Table 14.2.3.7 SHAPS Total Score – ANCOVA of Change from Baseline at Week 6 and Week 6 (LOCF) (ITT Population)

SAFETY DATA

Table 14.3.1.1 Adverse Event Overall Summary (Safety Population)

Table 14.3.1.2 Adverse Events, by System Organ Class and Preferred Term (Safety Population)

Table 14.3.1.3.1 Adverse Events, by System Organ Class and Preferred Term, by Maximum Severity (Safety Population)

Table 14.3.1.3.2 Adverse Events, by System Organ Class and Preferred Term, by Relationship to Study Medication (Safety Population)

Table 14.3.1.4 Adverse Events of Special Interest (AESI) by Preferred Term (Safety Population)

Table 14.3.1.5 Adverse Events Leading to Discontinuation from Study Treatment Period, by System Organ Class and Preferred Term (Safety Population)

Table 14.3.2.1 Serious Adverse Events, by System Organ Class and Preferred Term (Safety Population)

Table 14.3.4.1.1.1 Clinical Chemistry Laboratory Data in SI Units – Observed Value and Change from Baseline (Safety Population)

Table 14.3.4.1.1.2 Clinical Chemistry Laboratory Data in CV Units – Observed Value and Change from Baseline (Safety Population)

Table 14.3.4.1.2 Clinical Chemistry Laboratory Data – Shift from Baseline to Week 6 (Safety Population)

Table 14.3.4.2.1.1 Hematology Laboratory Data in SI Units – Observed Value and Change from Baseline (Safety Population)

Table 14.3.4.2.1 Hematology Laboratory Data in CV Units – Observed Value and Change from Baseline (Safety Population)

Table 14.3.4.2.2 Hematology Laboratory Data – Shift from Baseline to Week 6 (Safety Population)

Table 14.3.4.3.1 Urinalysis Laboratory Data – Observed Value and Change from Baseline (Safety Population)

Table 14.3.4.3.2 Urinalysis Laboratory Data (Categorical Values) – By Visit Summary (Safety Population)

Table 14.3.4.3.3 Urinalysis Laboratory Data – Shift from Baseline to Week 6 (Safety Population)

Table 14.3.4.4.1 Number and Percentage of Subjects with Potentially Clinically Significant (PCS) Post-Baseline Laboratory Values in SI Units (Safety Population)

Table 14.3.4.4.2 Number and Percentage of Subjects with Potentially Clinically Significant (PCS) Post-Baseline Laboratory Values in CV Units (Safety Population)

Table 14.3.5.1.1 Electrocardiogram (ECG) Parameters – Observed Value and Change from Baseline

(Safety Population)

Table 14.3.5.1.2 Electrocardiogram (ECG) Overall Assessment – By Visit Summary (Safety Population)

Table 14.3.5.1.3 Electrocardiogram (ECG) Overall Assessment – Shift from Baseline to Week 6 (Safety Population)

Table 14.3.5.1.4 Incidence of Potentially Clinically Significant (PCS) Electrocardiogram (ECG) Results (Safety Population)

Table 14.3.5.2.1 Vital Signs – Observed Value and Change from Baseline (Safety Population)

Table 14.3.5.2.2 BMI Categories – By Visit Summary (Safety Population)

Table 14.3.5.2.3 Number and Percentage of Subjects with Potentially Clinically Significant (PCS) Post-Baseline Vital Signs (Safety Population)

Table 14.3.5.2.4 Incidences of Orthostatic Hypotension and Orthostatic Tachycardia (Safety Population)

Table 14.3.5.3.1 C-SSRS – Any Suicidal Ideation, Suicidal Behavior, and Suicidality by Visit (Safety Population)

Table 14.3.5.3.2 C-SSRS – verall Post-baseline Subject Incidence of Treatment-Emergent Changes in Suicidality (Safety Population)

Table 14.3.5.4.1 AIMS Total Score – ANCOVA of Change from Baseline at Week 6 and Week 6 (LOCF) (Safety Population)

Table 14.3.5.4.2 AIMS Total Score – Shift from Baseline (Safety Population)

Table 14.3.5.4.3 AIMS Global Severity Score – Categorical Summary (Safety Population)

Table 14.3.5.4.4 AIMS Global Severity Score – Categorical Change from Baseline (Safety Population)

Table 14.3.5.5.1 BARS Total Score – ANCOVA of Change from Baseline at Week 6 and Week 6 (LOCF) (Safety Population)

Table 14.3.5.5.2 BARS Items – Categorical Summary (Safety Population)

Table 14.3.5.5.3 BARS Global Clinical Assessment – Categorical Change from Baseline (Safety Population)

Table 14.3.5.6.1 MSAS Total Score – ANCOVA of Change from Baseline at Week 6 and Week 6 (LOCF) (Safety Population)

Table 14.3.5.6.2 MSAS Total Score – Shift from Baseline (Safety Population)

Table 14.3.5.7.1 YMRS Total Score – Mixed Model for Repeated Measures Analysis of Change from Baseline over Time (Safety Population)

Table 14.3.5.7.2 YMRS Total Score - Treatment-Emergent Mania (Safety Population)

Table 14.3.5.8.1 Physician Withdrawal Checklist (PWC) Total Score – Observed value (Safety Population)

Table 14.3.5.8.2 Physician Withdrawal Checklist (PWC) Symptoms – Categorical Summary (Safety Population)

Table 14.3.5.9.1 SEP-4199 Plasma Concentrations (ng/mL) by Scheduled Time (PK Population)

Table 14.3.5.9.2 Plasma Prolactin Concentrations (ng/mL) by Scheduled Time (PK Population)

Figure 14.1 Time to Study Drug Discontinuation (ITT Population)

Figure 14.2.1.1 MADRS Total Score – LS Mean (95% CI) of Change from Baseline over Time Based on MMRM Estimates (ITT Population)

Figure 14.2.1.2 MADRS Total Score – LS Mean (95% CI) of Change from Baseline over Time Based on ANCOVA Estimates (ITT Population)

Figure 14.2.2.1 CGI-BP-S Depression Score – LS Mean (95% CI) of Change from Baseline over Time Based on MMRM Estimates (ITT Population)

Figure 14.2.2.2 CGI-BP-S Depression Score – LS Mean (95% CI) of Change from Baseline over Time Based on ANCOVA Estimates (ITT Population)

Listing 16.1.7 Randomization Scheme and Codes (Randomized Subjects)

Listing 16.2.1.1 Subject Disposition (Screen Failures)

Listing 16.2.1.2 Subject Disposition (Randomized Subjects)

Listing 16.2.1.3 Protocol Version and Study Medication Blinding Status (Randomized Subjects)

Listing 16.2.1.4 Analysis Population Assignment (Randomized Subjects)

Listing 16.2.2.1 Major and Critical Protocol Deviations (Randomized Subjects)

Listing 16.2.2.2 Inclusion Criteria (Randomized Subjects)

Listing 16.2.2.3 Exclusion Criteria (Randomized Subjects)

Listing 16.2.3.2 Subjects Excluded from Pharmacokinetic Analyses (Randomized Subjects)

Listing 16.2.4.1 Demographics and other Baseline Characteristics (Randomized Subjects)

Listing 16.2.4.2 Psychiatric History (Randomized Subjects)

Listing 16.2.4.3 Family Psychiatric History (Randomized Subjects)

Listing 16.2.4.4 Other Psychiatric disorder (Randomized Subjects)

Listing 16.2.4.5 Medical and Surgical History (Randomized Subjects)

Listing 16.2.4.6 Prior, Concomitant, and Post-treatment Medications (Randomized Subjects)

Listing 16.2.4.7 Timing of Last Dose of Benzodiazepines / Sedatives / Hypnotics (Randomized Subjects)

Listing 16.2.5.1 Drug Accountability (Randomized Subjects)

Listing 16.2.5.2 Randomized and Predominant Study Medication (Randomized Subjects)

Listing 16.2.6.1 Montgomery Asberg Depression Rating Scale (MADRS) (Randomized Subjects)

Listing 16.2.6.2 Clinical Global Impression – Severity Bipolar Version (CGI-BP-S) (Randomized Subjects)

Listing 16.2.6.3 Hamilton Rating Scale for Anxiety (HAM-A) (Randomized Subjects)

Listing 16.2.6.4 Quick Inventory of Depressive Symptomatology – Self-Report (QIDS-SR16) (Randomized Subjects)

Listing 16.2.6.5 Sheehan Disability Scale (SDS) (Randomized Subjects)

Listing 16.2.6.6 EuroQol 5D-5L (EQ-5D-5L) (Randomized Subjects)

Listing 16.2.6.7 Snaith-Hamilton Pleasure Scale (SHAPS) (Randomized Subjects)

Listing 16.2.7.1 Pre-treatment Events (Screen Failures)

Listing 16.2.7.2 Pre-Treatment Events (Randomized Subjects)

Listing 16.2.7.3 All Adverse Events (Safety Population)

Listing 16.2.7.4 Serious Adverse Events (Safety Population)

Listing 16.2.7.5 Adverse Events Leading to Discontinuation from Study Treatment Period (Safety Population)

Listing 16.2.7.6 Adverse Events Leading to Death (Safety Population)

Listing 16.2.8.1 Central Laboratory Data – Clinical Chemistry (Safety Population)

Listing 16.2.8.2 Central Laboratory Data – Hematology (Safety Population)

Listing 16.2.8.3 Central Laboratory Data – Urinalysis (Safety Population)

Listing 16.2.8.4 Central Laboratory Data – Other (Safety Population)

Listing 16.2.8.5 Central Laboratory Data – Urine Drug Screening (Safety Population)

Listing 16.2.8.6 Central Laboratory Data – Potentially Clinically Significant (PCS) Values (Safety Population)

Listing 16.2.8.7 Vital Signs (Safety Population)

Listing 16.2.8.8 Vital Signs – Potentially Clinically Significant (PCS) Values (Safety Population)

Listing 16.2.8.9 Electrocardiogram (ECG) Results (Safety Population)

Listing 16.2.8.10 Electrocardiogram (ECG) Results – Overall Assessment (Safety Population)

Listing 16.2.8.11 Columbia-Suicide Severity Rating Scale (C-SSRS) (Safety Population)

Listing 16.2.8.12 Pregnancy Test Results (Safety Population)

Listing 16.2.8.13 Serum Follicle Stimulating Hormone (FSH) Test Results (Safety Population)

Listing 16.2.8.14 Young Mania Rating Scale (YMRS) (Safety Population)

Listing 16.2.8.15 Abnormal Involuntary Movement Scale (AIMS) (Safety Population)

Listing 16.2.8.16 Barnes Akathisia Rating Scale (BARS) (Safety Population)

Listing 16.2.8.17 Modified Simpson-Angus Scale (MSAS) (Safety Population)

Listing 16.2.8.18 Physician Withdrawal Checklist (PWC) (safety Population)

Listing 16.2.8.19 Plasma SEP-4199 Concentrations (Safety Population)

Listing 16.2.8.20 Plasma Prolactin Concentration (Safety Population)

Listing 16.2.8.21 Plasma Concentrations for Subjects Excluded from Pharmacokinetic Analysis (Safety Population)

Listing 16.2.8.22 Visit Information (Randomized Subjects)

Listing 16.2.8.23 Comments Information (Randomized Subjects)

Statistical Appendix 16.1.9.2.1: Supportive Output for Table 14.2.1.1.2 MADRS Total Score – Mixed Model for Repeated Measures Analysis of Change from Baseline over Time (ITT Population)

Statistical Appendix 16.1.9.2.2: Supportive Output for Table 14.2.2.1 CGI-BP-S Score – Mixed Model for Repeated Measures Analysis of Change from Baseline over Time (ITT Population)

APPENDIX 10. SCHEDULE OF EVENTS

	Visit 1	Visit 2[a]	Visit 3	Visit 4	TC1[b]	Visit 5	TC2[b]	Visit 6[c]	Visit 7[d]
	Screening Washout[e]	Baseline	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6 EOT/ET	Follow-up
	Day -22 to -2	Day -1	Day 7 (± 2)	Day 14 (± 2)	Day 21 (± 2)	Day 28 (± 2)	Day 35 (± 2)	Day 42 (± 2)	7 (± 2) days after last dose
Obtain informed consent	X								
Obtain informed consent for duplicate subject check (where local regulations allow)	X								
Obtain informed consent for optional pharmacogenomic sampling, if applicable		X							
Lifetime Illness Characteristics Questionnaire[f]	X								
Montgomery-Asberg Depression Rating Scale, Self-rating version (MADRS-S)[f]	X								
Inclusion/Exclusion Criteria	X	X							
Duplicate Subject Check	X								
Pre-Baseline Eligibility Review[g]	X								
Demographics	X								
Medical History	X								
Psychiatric History[h]	X								
Family Psychiatric and Medical History	X								
Structured Clinical Interview for DSM-5-Clinical Trial Version (SCID-5-CT)[i]	X								
Physical Examination	X	X						X	X

	Visit 1	Visit 2[a]	Visit 3	Visit 4	TC1[b]	Visit 5	TC2[b]	Visit 6[c]	Visit 7[d]
	Screening Washout[e]	Baseline	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6 EOT/ET	Follow-up
	Day -22 to -2	Day -1	Day 7 (± 2)	Day 14 (± 2)	Day 21 (± 2)	Day 28 (± 2)	Day 35 (± 2)	Day 42 (± 2)	7 (± 2) days after last dose
Neurological Examination	X	X						X	X
Prior/Concomitant Medications	X	X	X	X	X	X	X	X	X
Subject Eligibility Check[j]		X							
Randomization		X							
Dispense Study Drug		X	X	X		X			
Study Drug Accountability[k]			X	X		X		X	
Schedule Next Visit[l]	X	X	X	X		X		X	
EFFICACY ASSESSMENTS									
Montgomery-Asberg Depression Rating Scale (MADRS)[m]	X	X	X	X		X		X	
Hamilton Anxiety Rating Scale (HAM-A)[n]		X	X	X		X		X	
Quick Inventory of Depressive Symptomatology-Self-Report (QIDS-SR16)		X						X	
Sheehan Disability Scale (SDS)		X						X	
EuroQol - 5 Dimension - 5 Level (EQ-5D-5L)		X						X	
Snaith-Hamilton Pleasure Scale (SHAPS)		X						X	
Clinical Global Impression-Bipolar Version-Severity of Illness (CGI-BP-S)[o]	X	X	X	X		X		X	

	Visit 1	Visit 2[a]	Visit 3	Visit 4	TC1[b]	Visit 5	TC2[b]	Visit 6[c]	Visit 7[d]
	Screening Washout[e]	Baseline	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6 EOT/ET	Follow-up
	Day -22 to -2	Day -1	Day 7 (± 2)	Day 14 (± 2)	Day 21 (± 2)	Day 28 (± 2)	Day 35 (± 2)	Day 42 (± 2)	7 (± 2) days after last dose
SAFETY ASSESSMENTS									
Vital Sign Measurements[p]	X	X	X	X		X		X	X
Height	X								
Weight and Body Mass Index[q]	X	X						X	X
Waist Circumference		X						X	
Pretreatment Event/ Adverse Events Monitoring[r]	X	X	X	X	X	X	X	X	X
12-Lead ECG	X	X	X	X		X		X	X
Serum Chemistry[s]	X	X		X				X	X
Hematology	X	X		X				X	X
Urinalysis	X	X		X				X	X
Serum Prolactin[t]	X	X		X				X	X
Hemoglobin A1c (HbA1c)	X	X						X	
Lipid Panel [s]	X	X						X	
Serum Insulin[s]	X	X						X	
High Sensitivity C-Reactive Protein (hs-CRP)	X	X						X	
Hepatitis B/C Panel	X								
Thyroid Panel	X	X						X	
Serum Follicle Stimulating Hormone (postmenopausal women or if menopause is suspected)	X								

	Visit 1	Visit 2[a]	Visit 3	Visit 4	TC1[b]	Visit 5	TC2[b]	Visit 6[c]	Visit 7[d]
	Screening Washout[e]	Baseline	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6 EOT/ET	Follow-up
	Day -22 to -2	Day -1	Day 7 (± 2)	Day 14 (± 2)	Day 21 (± 2)	Day 28 (± 2)	Day 35 (± 2)	Day 42 (± 2)	7 (± 2) days after last dose
Serum Pregnancy Test (females of childbearing potential)[u]	X								
Urine Pregnancy Test (females of childbearing potential)[u,v]		X		X				X	X
Urine Drug Screen[v]	X								
Rapid Urine Drug Test[v]		X		X				X	X
Young Mania Rating Scale (YMRS)[w]	X	X	X	X		X		X	X
Physician's Withdrawal Checklist (PWC)[x]								X	X
Columbia-Suicide Severity Rating Scale (C-SSRS)[y]	X	X	X	X		X		X	X
Abnormal Involuntary Movement Scale (AIMS)		X						X	
Barnes Akathisia Scale (BARS)		X						X	
Modified Simpson-Angus Scale (SAS)		X						X	
Blood Sampling for Plasma PK and Plasma Prolactin[t,z]	X	X		X				X	X
Optional Blood Sampling for Pharmacogenomic Testing[aa]		X							

Abbreviations: BMI = body mass index; eCRF = electronic case report form; DSM-5 = The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition; EOT = End of Treatment; ET = Early Termination; hCG = human chorionic gonadotropin; PK = pharmacokinetics; TC = telephone call.
Note: To ensure subject safety and data integrity, should circumstances warrant and with Sponsor approval, remote site/subject visits may be conducted.
There is an additional visit for Japan on Day 1 which checks study drug compliance and ECG.

[a] Visit 2 is defined as the Baseline Visit. Study drug will be dispensed at Visit 2/Day -1, and subjects will be instructed to take their first dose of study drug the following morning (Day 1).
[b] A telephone call at Day 21 and Day 35 to ascertain adverse events and concomitant medications since last visit.
[c] Visit 6 is the End of Treatment/Early Termination visit. Subjects who discontinue the study prior to Visit 6 will have all Visit 6 procedures performed at the time of discontinuation.
[d] Subjects who discontinue early from the study or complete the study and do not enter the extension study (SEP380-303) will have a safety Follow-up Visit (7 [± 2] days after their last dose of study drug).
[e] Screening assessments may occur over multiple days. Hospitalization during the Screening Period will not be allowed, except where required by local regulations or when determined to be clinically indicated based on the subject's psychiatric history and current psychiatric symptoms. Medical justification must be provided to the Medical Monitor and the Medical Monitor must approve the request for hospitalization. In such cases, a maximum of 7 days' hospitalization during washout of prior medications in the Screening Period will be allowed. An extension of the Screening Period by up to 7 days may be allowed, with prior approval from the Medical Monitor.
[f] Following Informed Consent, the Lifetime Illness Characteristics Questionnaire and the MADRS-S will be completed by the subject prior to other assessments at Screening. Subject must have a MADRS-S total score ≥ 22 at Screening.
[g] For the pre-Baseline review by the Sponsor Eligibility Committee, sites will be required to submit specific Screening information for CRO and Sponsor review, prior to proceeding to Baseline. Details are provided in Section 26, Appendix VII of the protocol.
[h] Includes a psychiatric history form that will include variables related to duration of illness, treatment response (eg, prior medications used to treat bipolar disorder) and other similar variables.
[i] The SCID-5-CT will be used to support the DSM-5 diagnosis and must be administered by a qualified rater at the site.
[j] Subject eligibility check in IXRS is to be performed prior to randomization, but after all scales are completed, to confirm Blinded Inclusion Criterion 12 is met.
[k] Clinical site staff will record the date and time of the 3 doses prior to the study visit in the source (all visits) and eCRF (Visit 4 [Day 14] and Visit 6 [Day 42] only), based on subject self-report.
[l] Instruct subject to record dosing date and time of the last 3 doses prior to the study visit on the blister pack wallet or other format and bring back all used/unused study drug and packaging to next visit.
[m] Subjects must have a MADRS total score ≥ 22 at Screening (Visit 1) and Baseline (Visit 2). Subjects who demonstrate a decrease (improvement) of ≥ 25% in MADRS total score from Screening (Visit 1) to Baseline (Visit 2) must be screen failed.
[n] The Structured Interview Guide for the HAM-A (SIGH-A) will be used for administration of the HAM-A.
[o] Subjects must have a CGI-BP-S depression score ≥ 4 at Screening (Visit 1) and Baseline (Visit 2).
[p] Blood pressure and pulse rate measurements will be taken in a supine and standing position. Blood pressure and pulse rate should first be taken with the subject in the supine position after resting for ≥ 5 minutes. Blood pressure and pulse rate will be taken again after standing for 2 to 4 minutes. The same arm should be used during each assessment of blood pressure and pulse rate throughout the study. Respiratory rate and temperature will also be measured.
[q] BMI will be calculated by clinical site staff at Screening (Visit 1) and recorded in the eCRF. At Baseline (Visits 2), Visit 6/EOT/ET (Day 42), and Visit 7 (Follow-up), BMI does not need to be calculated by the clinical site staff as it will be calculated in the analysis.
[r] Events occurring prior to first dose of study drug are programmatically identified as pretreatment events. Events occurring after first dose of study drug are programmatically identified as adverse events.
[s] Subjects are required to fast for at least 8 hours prior to sample collection for laboratory testing at Baseline (Visit 2) and Visit 6/EOT/ET (Day 42). Fasting for 8 hours prior to Screening (Visit 1) is also recommended to avoid potential retests.
[t] Prolactin levels (serum and plasma) will be masked for any assessment collected after the first study drug dose. Prolactin levels prior to the first study drug dose will be unmasked.
[u] For females of childbearing potential, any positive urine ß-hCG test should be confirmed by serum β-hCG.
[v] Unscheduled urine pregnancy (females of childbearing potential) and urine drug tests may be administered based on Investigator discretion. Point of care (POC) testing will be used for the urine pregnancy test and the rapid urine drug test. A urine drug screen should be submitted to the central lab for any positive rapid urine drug test throughout the study. Positive results for any of these assessments should be discussed with the Medical Monitor.
[w] To be eligible for enrollment, subjects must have a YMRS total score ≤ 12 at Screening (Visit 1) and Baseline (Visit 2).

[x] The PWC will be completed for early termination subjects and subjects who complete the study and do not rollover into Study SEP380-303; it will not be completed for subjects who rollover into Study SEP380-303.
[y] At the Screening visit (Visit 1), the "Baseline/Screening" version will be completed; for all subsequent visits, the "Since Last Visit" version of the C-SSRS will be administered.
[z] Blood sample for SEP-4199 population pharmacokinetic analysis for R- and S-enantiomers and/or plasma prolactin measurement will be collected at Screening (Visit 1), Baseline (Visit 2), Visit 4 (Day 14), Visit 6/EOT/ET (Day 42), and Visit 7/Follow-up (7 ± 2 days after last dose), with a record of the time of last 3 administered doses on the eCRF at Visit 4 (Day 14) and Visit 6/EOT/ET (Day 42). The blood sample will be collected at time of clinical safety laboratory test sample collection. Plasma concentrations of aramisulpride and esamisulpride and plasma prolactin levels will be measured. Remaining plasma from samples may also be used for the additional bioanalytical method development and/or characterization of putative metabolites of amisulpride and for other exploratory measurements, if needed.
[aa] Blood sampling for pharmacogenomic testing will be requested but is not required for participation in the study.