Official Title:	**A Multicenter, Randomized, Double-Blind, Parallel-Group, Placebo-Controlled Study Evaluating the Efficacy, Safety, and Pharmacokinetics of SAGE-217 in the Treatment of Adult Female Subjects With Severe Postpartum Depression**
NCT Number:	NCT02978326
Document Date:	SAP Version 1.0: 11 December 2018

Sage Therapeutics, Inc.

Statistical Analysis Plan

Methods

Protocol 217-PPD-201

A MULTICENTER, RANDOMIZED, DOUBLE-BLIND, PARALLEL-GROUP, PLACEBO-CONTROLLED STUDY EVALUATING THE EFFICACY, SAFETY, AND PHARMACOKINETICS OF SAGE-217 IN THE TREATMENT OF ADULT FEMALE SUBJECTS WITH SEVERE POSTPARTUM DEPRESSION

Author of SAP: [redacted], MStat

Version: Version 1.0

Date of SAP: 11 DECEMBER 2018

Sponsor
Sage Therapeutics, Inc.
215 First Street
Cambridge, Massachusetts 02142

SIGNATURE PAGE [redacted] OVALS

DocuSigned by:
Signer Name: [redacted]
Signing Reason: I approve this document
Signing Time: 18-Dec-2018 | 14:20 EST
882C53F2B5164EEAA9FA446FC310D506

18-Dec-2018 | 14:20:20 EST

[redacted], MStat — Date

Premier Research
387 Shuman Boulevard, STE 340W
Naperville, IL 60563
USA

DocuSigned by:
[redacted]
Signer Name [redacted]
Signing Reason: I approve this document
Signing Time: 18-Dec-2018 | 12:05 EST
BBA61679674E486C8EC496B3A80E1376

18-Dec-2018 | 12:05:27 EST

[redacted], Ph.D. — Date

Premier Research
387 Shuman Boulevard, STE 340W
Naperville, IL 60563
USA

DocuSigned by:
[redacted]
Signer Name: [redacted]
Signing Reason: I approve this document
Signing Time: 18-Dec-2018 | 13:27 EST
76E7B43E94C64D69A402F08ED43F79DB

18-Dec-2018 | 10:27:40 PST

[redacted], M.S — Date

Sage Therapeutics, Inc
215 First Street
Cambridge, MA 02142
USA

DocuSigned by:
[redacted]
Signer Name: [redacted]
Signing Reason: I approve this document
Signing Time: 20-Dec-2018 | 14:47 EST
94D2AB3B85B24D24889A649B36B6D73A

20-Dec-2018 | 14:47:53 EST

[redacted], MD, MBA — Date

Sage Therapeutics, Inc
215 First Street
Cambridge, MA 02142
USA

DocuSigned by:
[redacted]
Signer Name: [redacted]
Signing Reason: I approve this document
Signing Time: 18-Dec-2018 | 13:31 EST
D9A70EC8407D42CBB9016CDB10F3F851

18-Dec-2018 | 10:32:18 PST

[redacted] Ph.D Date
[redacted]
Sage Therapeutics, Inc
215 First Street
Cambridge, MA 02142
USA

DocuSigned by:
[redacted]
Signer Name: [redacted]
Signing Reason: I approve this document
Signing Time: 20-Dec-2018 | 10:02 EST
66D791A49C4941C1AE65924218CB7DDA

20-Dec-2018 | 10:02:21 EST

[redacted] Date
[redacted]
Sage Therapeutics, Inc
215 First Street
Cambridge, MA 02142
USA

1 TABLE OF CONTENTS

2 LIST OF ABBREVIATIONS

ABBREVIATION	DEFINITION OR DESCRIPTION
AE	adverse event
AR(1)	autoregressive
ATC	Anatomical Therapeutic Chemical
BAI	Beck Anxiety Index
BID	twice daily
[redacted]	
BMI	body mass index
BLQ	below the limit of quantification
%CV	percent coefficient of variation
CGI-I	Clinical Global Impression of Improvement
CGI-S	Clinical Global Impression of Severity
CI	confidence interval
C-SSRS	Columbia-Suicide Severity Rating Scale
CS	compound symmetry
CSR	clinical study report
ECG	electrocardiogram
eCRF	electronic case report form
[redacted]	
HAM-A	Hamilton Anxiety Rating Scale
HAM-D	Hamilton Rating Scale for Depression
[redacted]	
HEOR	Health economics outcomes
HIV	human immunodeficiency virus
ICF	informed consent form
MADRS	Montgomery-Åsberg Depression Rating Scale
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	mixed model for repeated measures
LLOQ	lower limit of quantitation
LS	least squares
[redacted]	
PPD	postpartum depression
PT	preferred term
QTcF	QT-interval for ECG corrected for heart rate (Fridericia)
SAE	serious adverse event
SAP	statistical analysis plan
SD	standard deviation
[redacted]	
SOC	System Organ Class
SSS	Stanford Sleepiness Scale
TEAE	treatment emergent adverse event
ULOQ	upper limit of quantitation
VC	variance components
WHO-DDE	World Health Organization Drug Dictionary Enhanced

3 INTRODUCTION

This statistical analysis plan (SAP) is for the final analysis of the study and is based on the approved clinical study protocol, dated 02 July 2018, Version 9.0, incorporating Amendment 8.

This SAP addresses the efficacy, safety, [redacted] objectives of the study and describes the planned efficacy, safety, [redacted] statistical analyses and data presentations.

The SAP described hereafter is an *a priori* plan of Sage study 217-PPD-201 and will be finalized and signed-off before database lock and treatment unblinding.

There are two parts to this study: Part A includes subjects enrolled under Protocol Version 2.0 (Amendment 1) or earlier, and includes an inpatient stay and treatment with SAGE-217 Oral Solution or matching placebo; in Part A, only one subject was enrolled and dosed before it was closed to enrollment. Therefore, data collected in Part A will only be listed. Part B includes subjects enrolled under Protocol Version 3.0 (Amendment 2) or later, and includes outpatient-only assessments and treatment with SAGE-217 Capsules or matching placebo. Starting with Protocol Version 5.0 (Amendment 4) (ie, PA4), the number of study visits was reduced from 9 visits (Days 2, 3, 4, 5, 6, 8, 15, 21 and 45) to 5 visits (Days 3, 8, 15, 21 and 45).

Data processing, tabulation of descriptive statistics, calculation of inferential statistics, and graphical representations will be performed primarily using SAS (release 9.4) for Windows.

[redacted]

All objectives and analyses described herein therefore are with respect to Part B, unless otherwise stated.

4 STUDY OBJECTIVES

4.1 Primary Objective

The primary efficacy objective of this study is to determine if treatment with SAGE-217 reduces depressive symptoms in subjects with severe postpartum depression (PPD) compared to placebo as assessed by the change from baseline in the Hamilton Rating Scale for Depression (HAM-D) total score at Day 15.

4.2 Secondary Objectives

The secondary efficacy objectives of this study are:

- To determine if treatment with SAGE-217 Capsules 30 mg QD reduces depressive symptoms in subjects with severe PPD compared to placebo as assessed by the change from baseline in the HAM-D total score for all other time points;
- To determine if treatment with SAGE-217 Capsules 30 mg QD reduces depressive symptoms compared to placebo as assessed by the
    - HAM-D response
    - HAM-D remission
    - Change from baseline to Day 15 in Montgomery-Åsberg Depression Rating Scale (MADRS) total score
    - Clinical Global Impression – Improvement (CGI-I) response
    - Changes from baseline in HAM-D subscales and individual item scores at Day 15 and all other time points

- To determine if treatment with SAGE-217 Capsules 30 mg QD reduces anxiety symptoms compared to placebo as assessed by changes from baseline in Hamilton Anxiety Rating Scale (HAM-A) total score at Day 15 and all other time points.

The safety objectives of this study are:

- To evaluate the safety and tolerability of SAGE-217 compared to placebo as assessed by the incidence of adverse events (AEs), vital sign measurements, clinical laboratory evaluations, electrocardiogram (ECG) parameters, and the Columbia Suicide Severity Rating Scale (C-SSRS).

[redacted]

5 STUDY ENDPOINTS

5.1 Efficacy Endpoints

5.1.1 Primary Efficacy Endpoint

The primary efficacy endpoint is the change from baseline in the HAM-D total score at the end of the treatment period (Day 15).

5.1.2 Secondary Efficacy Endpoints

The secondary efficacy endpoints include the following:

- Change from baseline in the HAM-D total score at all other time points;
- HAM-D response, defined as a 50% or greater reduction from baseline in HAM-D total score;
- HAM-D remission, defined as a HAM-D total score of ≤ 7;
- Change from baseline in MADRS total score and individual items scores at Day 15 and other time points;
- CGI-I response, defined as "very much improved" or "much improved";
- Change from baseline in HAM-A total score at Day 15 and other time points; and

- Change from baseline in HAM-D subscales and individual item scores at Day 15 and other time points.

[redacted]

5.2 Safety Endpoints

The safety and tolerability endpoints include:

- Frequency and severity of AEs; and
- Observed values and changes from baseline in vital signs measurements, clinical laboratory data, ECG parameters, and incidences of suicidal ideation or behaviors using the C-SSRS.

[redacted]

6 STUDY DESIGN

6.1 Overall Design

This is a multicenter, randomized, double-blind, parallel-group, placebo-controlled study of the efficacy, safety, and PK of SAGE-217 in adult female subjects diagnosed with severe PPD.

Subjects will receive SAGE-217 Capsules 30 mg QD or placebo. Those subjects who cannot tolerate 30 mg QD will receive 20 mg QD for the remainder of the Treatment Period. Subjects who experience intolerable AEs at the 20 mg QD dose level may be discontinued from study treatment at the discretion of the Investigator.

The study will consist of an up to 28-day Screening Period (Day -7 or -28, to -1), 14-day Treatment Period (Day 1 to Day 14), and Follow-up Period (through Day 45). The screening assessment will be conducted on an outpatient basis. During the study Treatment Period, which is a total of 14 days, and Follow-up Period assessments will be conducted on an outpatient basis.

During the Screening Period, after signing the informed consent form (ICF), subjects will be assessed for study eligibility and the severity of each subject's PPD will be evaluated using HAM-D. Eligible subjects will be admitted to the clinical study unit on Day 1.

If applicable, standard of care data collected prior to obtaining informed consent may also be included as screening data, if appropriate, such as laboratory tests, ECG, physical examination, and vital signs conducted within the preceding 48 hours, as long as the requirement for the screening assessment to be

collected retrospectively is met in full. If applicable, to ensure protocol compliance, any standard of care data eligible for inclusion as screening data must include the precise nature and timing of data collection.

Eligible subjects will be randomized to active study drug or placebo on a 1:1 basis. The end of the Screening Period coincides with the beginning of the Treatment Period.

Efficacy and safety assessments will be performed periodically during the study, [redacted]

Subjects will be monitored for safety during the Treatment and Follow-up Periods (through Study Day 45 [±3 days]) including monitoring for AEs/serious AEs (SAEs), routine clinical laboratory assessments, physical examination, vital signs, and ECG.

6.2 Sample Size and Power

Assuming a 2-sided test at an alpha level of 0.05, a sample size of approximately 65 evaluable subjects per treatment group would provide 90% power to detect a placebo-adjusted treatment difference of approximately 4 points in the primary endpoint, change from baseline in HAM-D total score at Day 15 assuming standard deviation (SD) of 7 points.

Assuming a 10% dropout and a 1:1 randomization ratio, approximately 72 randomized subjects per treatment group will be required to obtain 130 evaluable subjects. Evaluable subjects are defined as those randomized subjects who received study drug and have a valid baseline and at least 1 post-baseline HAM-D assessment. Additional subjects may be randomized if the dropout rate is higher than 10%.

6.3 Randomization

Subjects who meet the entrance criteria will be randomly assigned in a 1:1 ratio to receive SAGE-217 or placebo according to a randomization schedule. Once it has been determined that a subject meets eligibility criteria, the subject will be sequentially assigned a randomization number from the randomization schedule provided to the unblinded pharmacist and/or designated pharmacy staff (prior to protocol 6.0) or via interactive response technology (IRT) system (Protocol amendment 6.0 and afterwards). Randomization numbers will consist of a 4-digit number starting with a 3 (eg, "[redacted]") for paper (manual) randomization, or starting with a 4 (eg, "[redacted]") for IRT randomization occurring for subjects enrolled under Amendment 6.0 or afterwards.

Subject randomization (including study part, subject identifier, randomization number, and treatment assigned) will be provided in a listing.

6.4 Blinding and Unblinding

Subjects, clinicians, and the study team will be blinded to treatment allocation.

Randomization schedules will be generated by an independent statistician. The allocation to treatment groups (SAGE-217 or placebo) will be based on the randomization schedule. The randomization schedules will be kept strictly confidential, accessible only to authorized personnel until the time of unblinding.

In the event of a medical emergency, the pharmacist and/or designated pharmacy staff may reveal actual study drug contents to the Investigator, who should also alert Sage of the emergency. In all cases where the study drug allocation for a subject is unblinded, pertinent information (including the reason for unblinding) must be documented in the subject's records and on the electronic case report form (eCRF). If the subject or study center personnel (other than pharmacist and/or designated pharmacy staff) have been unblinded, the subject will be terminated from the study.

7 MODIFICATIONS

7.1 Modifications to the Approved Clinical Study Protocol

The following changes from the clinical study protocol (version 9.0, dated July 2, 2018) have been made in this SAP:

1. [redacted]
2. The protocol (Section 14.6.2.1) defines the upper limit for a TEAE as 7 days after last dose of study drug, whereas Section 13.4 of the protocol defines TEAEs as any AE occurring after first administration of study drug. Per Sage's request, this SAP follows Section 13.4 of the protocol in defining TEAEs.
3. The protocol does not have MADRS response and remission as secondary endpoints, but this was requested to be analyzed. Additionally, subgroup analyses of selected efficacy endpoints were requested by the Sponsor.
4. The protocol has the All Randomized Set as a data analysis set, used to analyze subject disposition, demographics, and baseline characteristics summaries; however, per Sage's request, subject disposition includes all enrolled subjects, and demographics and baseline characteristics will be analyzed using subjects in the Safety Set and the Efficacy Set.
5. Only one subject was enrolled and dosed in Part A before it was closed to enrollment. Therefore, data collected in Part A will only be listed.
6. Changes to the original and subsequent sample size estimates: Each amendment in which there was a change to the desired study power and placebo-adjusted treatment difference for this study was guided by clinical information external from this study. This information includes (1) comments from the Agency during the brexanolone PPD Program Breakthrough Therapy (BTD) designation (BTD) meeting, (2) study results from the brexanolone program (547-PPD-202B and 202C) and (3) study results from the SAGE-217 MDD program (217-MDD-201B). The original sample size calculation used 10% alpha level 80% power with treatment difference and SD of 10 points each (based on observed data from SAGE-547-PPD-202A). Following discussions with the FDA related to BTD for brexanolone, comments were incorporated into the SAGE-217 PPD study to use alpha 5% which led to a decrease in the power to 72%. Consequently, following the availability of study result from SAGE-217-MDD-201 Part B that showed 7.5 points (smaller than 10 points) placebo-adjusted treatment difference, the SAGE-217 PPD-201 study sample size was increased to maintain study power at the pre-specified 80%. Finally, to position the study as a potential pivotal study, Sage took the opportunity to increase study power to 90% based on a placebo-adjusted treatment difference of 4 points, SD = 7 points (per pooled SAGE-547-PPD-202A, 202B, 202C 90 mg dose data). This was described in Amendment 6 of the protocol and SAP (Section 6.2).
7. Sensitivity analysis was added in case more than 10% of subjects are missing HAM-D scores at Day 15.
8. Per protocol analysis is added.
9. [redacted]

7.2 Modifications to the Approved Statistical Analysis Plan

This is the first version of the SAP.

7.3 Modifications to the Approved DMC Charter

Not applicable (No DMC for this study).

8 ANALYSIS SETS

Subjects included in the below analysis sets (and reason for exclusion, if applicable) will be provided in a listing.

8.1 Efficacy Set

The Efficacy Set will consist of all subjects who are administered study drug and have a valid baseline and at least 1 post-baseline efficacy assessment. The Efficacy Set will be used to analyze all efficacy data and demographics and baseline characteristics data according to randomized treatment group.

8.2 Safety Set

The Safety Set will consist of all subjects who are administered study drug. This analysis set will be used to provide descriptive summaries of safety data according to treatment received. If the subject took any dose of SAGE-217 during the study, she will be considered under SAGE-217 for treatment received. Historical data (eg, demographics and baseline characteristics, prior medications, and psychiatric history) will be summarized by treatment received.

8.3 Per Protocol Set

The Per Protocol (PP) set will include subjects in the Efficacy set and without significant protocol violations or deviations, which in opinion of the study team, could affect efficacy. This review of major protocol deviations will be completed, and the decision on whether the deviation affects efficacy will be documented before database unblinding. Subjects will be classified according to randomized treatment. This analysis set will be used for sensitivity analyses of the primary endpoint when at least 5% of subjects had major protocol deviations.

9 STATISTICAL ANALYSIS

9.1 General Considerations

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

Categorical (qualitative) variables will be summarized using the number and proportion of each possible value. In general, the denominator for the percentage calculation will be based upon the total number of subjects in the study population for the treatment group, unless otherwise specified.

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

Percentages will be presented to 1 decimal place unless otherwise specified. Assessments done on unscheduled visits or on unscheduled parameters will not be summarized but will be listed unless otherwise specified.

Treatment received or randomized treatment is defined as SAGE-217 or placebo. If the subject took any dose of SAGE-217 during the study, she will be considered under SAGE-217 for treatment received. All background summaries will be presented by randomized treatment and overall subjects. All safety summaries will be presented by treatment received. All efficacy summaries will be presented by treatment group. [redacted]

In the summary table and figures, the visit presented will follow the most recent protocol amendment. Any assessment under visits designated in the previous protocol will be included in the listing.

All final, planned analyses identified in the protocol and in this SAP will be performed after all relevant study data have been processed and integrated into the analysis database, analysis populations have been finalized, and the database has been locked. Any post-hoc, exploratory analysis performed to support planned study analyses, which are not identified in this SAP, will be documented and reported in Section 9.8 of the CSR. Any results from these unplanned analyses (post-hoc) will also be clearly identified in the text of the CSR.

All statistical tests will be performed as 2-sided tests with a significance level (alpha) of 0.05.

All collected data for both study parts will be presented in listings and will be sorted by study part and subject.

If partial dates occur, the convention for replacing missing dates for the purposes of calculating derived variables will be as follows:

The following conventions will be used for incomplete or missing medication dates to determine whether the medications were concomitant:

Concomitant Medication

If the start date (or end date) of medication is completely missing (ie, in which the day, month, and year are all unknown) or only the day is known, then the start date (or end date) will not be imputed and the medication will be considered to be concomitant.

For the partial start date of medication:

- If the year is present and the month and day are missing, then the month and day will be set to January 1.
- If the year and day are present and the month is missing, then the month will be set to January.
- If the year and month are present and the day is missing, then the day will be set to the 1st day of month.
- If the imputed start date of medication is after the non-imputed end date of medication, then the start date will be set to the end date of medication.

For the partial end date of medication:

- If the year is present and the month and day are missing, then the month and day will be set to December 31.
- If the year and day are present and the month is missing, then the month will be set to December. If the year and month are present and the day is missing, then the day will be set to the last day of the month.

Adverse Event

If AE start dates are completely missing or partially missing, the following date imputation rules will be applied for the determination of treatment-emergence:

If an AE onset date is completely missing (ie, in which the day, month, and year are all unknown), then the AE onset date and time will be set to the date and time of initiation of the treatment.

For partial AE start dates:

- When the year is known, but the month and day is unknown, then:
  - If the year matches the year of first dose date and the end date (if present) is after first dose date, or AE is ongoing, then impute as the month and day of the first dose date.
  - If the year of AE onset < year of initiation of the treatment, then the month and day will be set to December 31st.
  - If the year of AE onset > the year of initiation of treatment, then the month and day will be set to January 1st.

- When the year and day are present and the month is missing:
  - If the year of AE onset = the year of initiation of the treatment, then the month will be set to the month of initiation of the treatment.
  - If the year of AE onset < the year of initiation of the treatment, then the month will be set to December.
  - If the year of AE onset > the year of initiation of the treatment, then the month will be set to January.

- If the year and month are known, but the day is unknown, then:
  - If the year of AE onset = the year of initiation of the treatment and:
    - the month of AE onset = the month of initiation of the treatment, then the day will be set to the day of initiation of the treatment.
    - the month of AE onset < the month of initiation of the treatment, then the day will be set to the last day of month.
    - if the month of AE onset > the month of initiation of the treatment, then the day will be set to the 1st day of month.
  - If the year of AE onset < the year of initiation of the treatment, then the day will be set to the last day of month.
  - If the year of AE onset > the year of initiation of the treatment, then the day will be set to the 1st day of month.

If the imputed AE onset date is after the AE stop date, then the onset date will be set to the stop date.

If an AE has a missing severity or relationship to study drug, no imputations will be performed. No missing data will be imputed unless otherwise specified.

In general, for quantitative laboratory values reported as '<X' or '≤X', the lower limit of quantitation (LLOQ) will be used for analysis (ie, a value of X will be used in the analysis for lab values reported as '<X' or '≤X'). Similarly, for quantitative laboratory values reported as '>X' or '≥X', the upper limit of quantitation (ULOQ) will be used for analysis (ie, a value of X will be used in the analysis for lab values reported as '>X' or '≥X').

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

The last measurement before the start of blinded study drug administration will be used as the baseline observation for all calculations of change from baseline.

9.2 Background Characteristics

9.2.1 Subject Disposition

All subjects who sign the ICF will be accounted for in this study. The number of subjects enrolled (ie, screened) will be summarized by overall subjects. The number of subjects randomized and who received study drug, treatment group, and overall subjects. The number and percentage of subjects who completed the study, who withdrew from the study early (as well as the reason for withdrawal), and in each analysis population will be summarized by part, treatment group and overall subjects.

All disposition information will be included in a listing.

9.2.2 Demographics and Baseline Characteristics

Demographics, such as age, age categories (18-24, 25-45), child-bearing potential, race, and ethnicity, and baseline characteristics such as height, weight, body mass index (BMI), BMI Categories (≤18.4, 18.5-24.9, 25-29.9, ≥30 kg/m^2), Baseline antidepressant use (Yes, No), family history of PPD (Yes, No) and onset of PPD (3^{rd} trimester, within 4 weeks of delivery) will be summarized by treatment group and overall subjects.

Subjects who marked more than one race will be summarized in a "More than One Race" category. The demographic and baseline characteristics table will be generated for subjects in the Safety Set and Efficacy Set.

Medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) coding system (version 19.1 or higher). Frequencies and percentages of subjects with a previous depression or anxiety diagnosis will have the following summarized by treatment group and overall subjects for subjects in the Efficacy Set: type of depression or anxiety diagnosis, number of episodes of PPD (1, 2, 3, >3), number of hospitalizations (1, 2, 3, >3) as a result of depression or anxiety, and number of subjects with a family member with a history of psychiatric conditions.

Hepatitis, human immunodeficiency virus (HIV), drug and alcohol, and pregnancy screening results will be listed, but not summarized as they are considered part of the inclusion/exclusion criteria.

Medical history and psychiatric family history will be listed.

9.2.3 Prior and Concomitant Medications

Concomitant medications will be coded using World Health Organization Drug Dictionary Enhanced (WHO-DDE) (September 2016).

Frequencies and percentages of medications used in the study will be summarized as follows:

- Prior medication: medication with a start date before the first dose of study drug (including medications with a missing start date but a non-missing end date in this time period).
- Concomitant medication: medication with a start date on or after the first dose of study drug, or those with a start date before the first dose of study drug that are ongoing or with a stop date on or after the first dose of study drug. If medication dates are incomplete and it is not clear whether the medication was concomitant, it will be assumed to be concomitant.

Details of prior and concomitant medications will be listed by study part, treatment group, subject, start date, and verbatim term. All medications will be summarized by Anatomic Therapeutic Chemical (ATC) class and preferred term (PT). Furthermore, all medications will be summarized by treatment group and overall subjects in. Subjects will be counted once per ATC and PT.

Antipsychotic and antidepressant medications (ie, all medications belonging to the ATC 3 codes N05A and N06A respectively) used at treatment start (ie, a prior medication that is ongoing at treatment start) will be summarized and listed in a similar manner as prior and concomitant medications.

Any new antidepressant medication or dose reduced/increased will be identified and summarized by study period (treatment period, and follow-up period) and treatment group.

Different flags identifying antidepressants and antipsychotics will be included in the prior and concomitant medications listing.

Medication summaries will be based on the Efficacy Set.

9.2.4 Study Drug Exposure and Compliance

For all of the below formulas, *i* = individual subject for the entire treatment period (unless otherwise specified).

- Total drug exposure in mg for SAGE-217 will be calculated as:

$$\sum_i Actual\ Doses\ Taken\ \ (mg)$$

- Total drug exposure for placebo subjects is zero.
- Average daily dose in mg/day will be calculated as:

$$\frac{\sum_i Actual\ Doses\ Taken\ \ (mg)}{\sum_i Days\ on\ Study\ Drug} \times 100$$

- Total days of exposure will be calculated as:

$$\sum_i Days\ on\ Study\ Drug$$

- Percent of the planned exposure received will be calculated as:

$$\frac{\sum_i Actual\ Doses\ Exposure\ \ (mg)}{\sum_i Planned\ Exposure\ (mg)} \times 100$$

where $\sum_i Days\ on\ Study\ Drug$ = last dose date – first dose date +1. For subjects who complete the treatment, planned exposure is 14 days of treatment planned, times X mg for subjects randomized to SAGE-217, where X=30 for subjects randomized to 30mg, or X=20 for if subjects had dose reduction to 20mg. For subjects who discontinued the treatment early, the planned exposure is calculated as the number of doses planned up to the point of discontinuation. $\sum_i Actual\ Doses\ Exposue\ \ (mg)$ = sum of all actual doses taken on

each day. If a subject only took one capsule on a given day, the actual dose is assumed to be 30mg if the planned dose is 30mg on that day or 10 mg if the planned dose is 20 mg on that day.

- Study drug compliance (%) is defined as the number of capsules taken, divided by the number of capsules planned to be taken, times 100. Planned number of capsules taken is 28 for 14 days of two-per-day treatment for subjects who completes the treatment period. For subjects who discontinue the treatment early, the planned number of capsules is (Last dose date – First dose date + 1) x 2. If a kit is not returned, compliance will be left missing.

Total drug exposure (mg), average daily dose (mg/day), total days of exposure, Percent of the planned exposure received and compliance (%) will be summarized descriptively by treatment group and overall.

Study drug dosing and compliance information (including unplanned dose adjustments) will be listed. Any other study drug noncompliance such as missing visits, interruptions in the schedule of administration, and nonpermitted medications will be listed in the protocol deviations listing.

9.3 Efficacy Analysis

The primary efficacy variable for this study is the change from baseline in HAM-D total score at the end of the Treatment Period (Day 15).

The secondary efficacy variables for this study include:

- Change from baseline in HAM-D total score at all time points other than Day 15;
- HAM-D response defined as 50% or greater reduction from baseline in HAM-D total score;
- HAM-D remission defined as a HAM-D total score of ≤7;
- Change from baseline in MADRS total score at Day 15 and other time points;
- Change from baseline in CGI-S total score at Day 15 and other time points;
- CGI-I response defined as “very much improved” or “much improved”;
- Change from baseline in HAM-A total score at day 15 and other time points;
- Changes from baseline in HAM-D subscales and individual item scores at Day 15 and other time points

[redacted]

Efficacy analyses will be based on Efficacy Set using randomized treatment as treatment group.

For the windowing of visit in the efficacy analysis, please refer to section 9.3.4.

9.3.1 Analysis of Primary Efficacy Variable

9.3.1.1 Description of Primary Efficacy Variable

The primary outcome measurement is the change from baseline in 17-item HAM-D total score at the end of the Treatment Period (Day 14), based on the Day 15 morning assessment.

The 17-item HAM-D will be used to rate the severity of depression in subjects who are already diagnosed as depressed (Hamilton 1960). The 17-item HAM-D comprises individual ratings related to the following symptoms: depressed mood (sadness, hopeless, helpless, worthless), feelings of guilt, suicide, insomnia (early, middle, late), work and activities, retardation (slowness of thought and speech; impaired ability to concentrate; decreased motor activity), agitation, anxiety (psychic and somatic), somatic symptoms (gastrointestinal and general), genital symptoms, hypochondriasis, loss of weight, and insight. The HAM-D assessments are to be completed within ±30 minutes of the scheduled time point, but before starting dosing on Day 1. Every effort should be made for the same rater to perform all HAM-D assessments for an individual subject. If more than 3 individual items are missing, the HAM-D total score will not be calculated and will be left as missing. If less than or equal to 3 individual item scores are missing, the missing item scores will be imputed by the maximum possible value of the item or the mean of all other available item scores, whichever is smaller. Imputed individual scores will be rounded to the nearest integer.

Assessments of insomnia items during the Treatment Period will be rated relative to the "last night's sleep". The HAM-D total score will be calculated as the sum of the 17 individual item scores.

The HAM-D subscales are Core, Anxiety, Bech-6, and Meier. Refer to following table for details on HAM-D Subscales and Items for Calculating Each Subscale score:

HAM-D Subscales	**Items**	**Calculation**
Core	Depressed mood Feeling of guilt Suicide Work and activities Retardation	Sum of the 5-item responses/20 x 100. If more than one item responses are missing or HAM-D total score is missing, leave as missing; otherwise, use the imputed item score used to calculate HAM-D total score to calculate the subscale.
Anxiety	Anxiety psychic Anxiety somatic Somatic symptoms gastrointestinal Somatic symptoms general Hypochondriasis Loss weight	Sum of the 6-item responses/18 x 100. If more than one item responses are missing or HAM-D total score is missing, leave as missing; otherwise, use the imputed item score used to calculate HAM-D total score to calculate the subscale.
Bech-6	Depressed mood Feeling of guilt Work and activities Retardation Anxiety psychic Somatic symptoms general	Sum of the 6-item responses/22 x 100. If more than one item responses are missing or HAM-D total score is missing, leave as missing; otherwise, use the imputed item score used to calculate HAM-D total score to calculate the subscale.
Meier	Depressed mood Feeling of guilt Work and activities Retardation Agitation Anxiety psychic	Sum of the 6-item responses/24 x 100. If more than one item responses are missing or HAM-D total score is missing, leave as missing; otherwise, use the imputed item score used to calculate HAM-D total score to calculate the subscale.

Before PA4, the HAM-D will be collected in the clinic at Screening; before dosing on Day 1; 12 hours ± 1 hour following the evening dose on Days 2, 3, 4, 5, 6, 8, and 15; and in the morning on Days 21 and 45. On or after PA4, the HAM-D will be collected at the scheduled clinic visits at Screening; before dosing on Day 1; 12 hours (±1 hour) following the previous evening's dose on Days 3, 8, and 15; and in the morning on Days 21 and 45.

9.3.1.2 Visit Windows for the Primary Variable

For efficacy analyses, unscheduled measurements will only be included if a scheduled measurement is not available and the unscheduled measurement falls on the same study day. Visit windows are defined in Section 9.3.4.

9.3.1.3 Primary Analysis

Observed values and change/percent change from baseline to the Day 15 morning assessment (Day 15 – Baseline, and all other time points) in HAM-D total score and individual item scores will be summarized by treatment group. Lower scores postdose indicate improvement in depressive symptoms.

The difference between treatment groups in change from baseline to Day 15 in HAM-D total score will be evaluated by an MMRM with treatment, baseline HAM-D total score, baseline antidepressant use, assessment time point, and time point-by-treatment interaction as fixed effects. All post-baseline time points that are collected subjects both before and after release of PA4 (ie, Days 3, 8, 15, 21, and 45) will be included in the model; however, the primary comparison will be between SAGE-217 and placebo at the Day 15 time point.

In comparing the visit value to baseline within each treatment group, least squares (LS) mean change from baseline will be presented along with associated 95% confidence interval (CI). In comparing treatments (SAGE-217 minus placebo), LS mean differences in change from baseline will be presented along with associated 95% CI and pairwise treatment *P* value. An unstructured covariance structure will be used to model the within-subject errors. The Kenward-Roger approximation will be used to estimate denominator degrees of freedom.

If there is a convergence issue with the unstructured covariance model, Toeplitz, compound symmetry or Autoregressive (1) (AR[1]) covariance structure will be used, following this sequence until convergence is achieved. If the model still does not converge with AR(1) structure, no results will be reported. When the covariance structure is not UN, the sandwich estimator for the variance-covariance matrix will be derived, using the EMPIRICAL option in the PROC MIXED statement in SAS.

The LS mean change from baseline (and 95% CI of the mean) in HAM-D total score based on MMRM results will be plotted over time by treatment group.

9.3.1.4 Supportive Analyses

The change from baseline in HAM-D total scores will be analyzed by MMRM methods similar to the primary endpoint, as described in Section 9.3.1.3, with the addition of pooled center in the model. Pooling of centers is described in Section 9.3.5.

9.3.1.5 Multiplicity Adjustment

No adjustments for multiplicity will be performed for this Phase 3 trial. Only the primary endpoint analysis will be treated as a formal hypothesis test, and the secondary endpoints will be considered as supportive of the primary endpoint.

9.3.1.6 Sensitivity Analysis

1) Sensitivity analyses will be carried out to investigate the impact of missing data if more than 10% of subjects are missing primary endpoint assessment (ie, HAM-D total score at Day 15).

A dropout reason based imputation will be used. An analysis of covariance (ANCOVA) model, including treatment, baseline antidepressant use, and baseline HAM-D total score, will be used to assess the treatment difference in change from baseline in HAM-D total score at Day 15 (MI imputed data). The missing change from baseline in HAM-D total score at Day 15 will be imputed using multiple imputation.

Imputation distribution:
The imputation distribution for the missing change from baseline in HAM-D total score at Day 15 will be a normal distribution. All treated subjects will be classified as non-missing category, missing category 1, or missing category 2, based on the following rules:

- Non-missing category: Subject with non-missing change from baseline in HAM-D total score at Day 15.
- Missing category 1: Subject discontinued due to adverse events, physician decision, protocol violation or other, and is missing change from baseline in HAM-D total score at Day 15.
- Missing category 2: Subject discontinued due to pregnancy, study terminated by sponsor, or withdrawal by subject, and is missing change from baseline in HAM-D total score at Day 15, or subject completed study but is missing change from baseline in HAM-D total score at Day 15.

Imputation algorithm:
Missing values of change from baseline in HAM-D total score at Day 15 will be imputed separately within each treatment group using the following missing reason based algorithm:

- **Missing category 1:** randomly draw a sample from the normal distribution $N(\mu_{75}, \sigma^2)$, where μ_{75} is the 75th percentile of the non-missing change from baseline in HAM-D total score and σ^2 is the sample variance estimated using the non-missing change at Day 15. This represents a conservative approach since higher values of change from baseline represents worse outcome.
- **Missing category 2:** randomly draw a sample from the normal distribution $N(\mu, \sigma^2)$, where μ is the mean of the non-missing change from baseline in HAM-D total score at Day 15 and σ^2 is the sample variance estimated using the non-missing change at Day 15.

Analysis model:
The complete MI method is described below:

- Impute missing values using the normal distribution specified in the above algorithm to form a complete dataset (imputed dataset). After imputation, all Efficacy Set subjects will have non-missing change from baseline in HAM-D total score at Day 15.
- Repeat the process K (K=5) times, using the procedure described above to form K imputed complete datasets.
- Fit the ANCOVA model including baseline antidepressant use, and baseline HAM-D total score, to each imputed dataset, to estimate the treatment effect and its variance at Day 15.
- Combine the results from the K imputed datasets using the SAS procedure MIANALYZE, to derive the MI estimator.

We fit the analysis model (ANCOVA model specified before) to the kth completed dataset, denoting the estimate of the treatment difference θ by θ_k from the kth completed dataset, and denoting the corresponding estimate of the variance V_k. The MI estimator of θ, $\tilde{\theta}_{MI}$, is the

average of the K individual estimators:

$$\tilde{\theta}_{MI} = \frac{1}{K}\sum_{k=1}^{K}\theta_k$$

The estimated variance of $\tilde{\theta}_{MI}$ is a combination of the between- and within-imputation variability as follows:

$$V_{MI} = W + \left(1 + \frac{1}{K}\right)B$$

where $\mathrm{W} = \frac{1}{K}\sum_{k=1}^{K} V_K$ is the within-imputation variability and $\mathrm{B} = \frac{1}{K-1}\sum_{k=1}^{K}\left(\theta_k - \tilde{\theta}_{MI}\right)^2$ is the between-imputation variance.

It has been shown that the statistic

$$\mathrm{T} = \frac{\tilde{\theta}_{MI} - \theta}{\sqrt{V_{MI}}}$$

has an approximate t_V distribution where $\mathrm{V} = (\mathrm{K}-1)\left(1 + \frac{W}{B}\right)^2$.

2) The following sensitivity analysis will be performed to account for the possible scenario that the data is not normally distributed.

Stratified Wilcoxon rank-sum test on change from baseline in the HAM-D total score at Day 15, stratified by antidepressant use, will be used as a sensitivity analysis. If necessary (e.g., serious violation of the primary analysis model assumptions on normality, as assessed by using QQ plot on the residual values from the primary analysis for HAM-D), this will be considered the primary analysis.

See sample SAS code for stratified Wilcoxon rank-sum test using PROC FREQ in section 12.3.

9.3.1.7 Subgroup Analysis for Primary Endpoint

The MMRM analysis for change from baseline in HAM-D total score analysis will be repeated for the following subgroups:

- Race category: White, Black or African American, and Other
- Age category: 18-24, 25-45
- Baseline antidepressant use: Yes, No
- BMI category: ≤18.4, 18.5-24.9, 25-29.9, ≥30 kg/m^2
- Onset of PPD: 3rd trimester, within 4 weeks of delivery
- Family history of PPD: Yes, No

In the MMRM model for subgroup analysis, antidepressant use will be excluded from the model due to the small number of subjects in antidepressant use group in each subgroup.

Observed values and change/percent change from baseline to the Day 15 morning assessment (Day 15 – Baseline, and all other time points) in HAM-D total score will be summarized by treatment group and subgroup.

In addition, the LS mean change from baseline (and 95% CI of the mean) in HAM-D total score at Day 15 based on MMRM results will be plotted by treatment group and subgroup.

9.3.2 Analysis of Secondary Efficacy Variables

All secondary analyses will be conducted using the Efficacy Set.

9.3.2.1 Hamilton Rating Scale for Depression (HAM-D)

Description of HAM-D Response and Remission

The HAM-D response is defined as a 50% or greater reduction from baseline in HAM-D total score (ie, percent change from baseline ≤-50%); HAM-D remission is defined as a HAM-D total score of ≤7.

Analysis of HAM-D Response and Remission

Number and percentage of subjects meeting the criteria for HAM-D response and HAM-D remission will be summarized by time point and treatment group.

The number of subjects meeting HAM-D response and HAM-D remission criteria at each time point will be analyzed using GEE for binary response. The model will include terms for treatment, baseline HAM-D total score, baseline antidepressant use, assessment time point, and time point-by-treatment interaction. In comparing treatments (SAGE-217 minus placebo), model-based point estimates (ie, adjusted odds ratios) will be presented along with associated 95% CI and pairwise treatment *P* value.

An unstructured covariance structure will be used to model the within-subject errors. In case of convergence issues, other covariance structures will be used including Toeplitz, compound symmetry or Autoregressive (1) (AR[1]) covariance structure will be used, following this sequence until convergence is achieved. If the model still does not converge with AR(1) structure, no results will be reported.

In addition, as a supportive analysis, a logistic regression model will be used separately for each visit, which will include the response variables above as the dependent variable, baseline antidepressant use, treatment, and baseline HAM-D total score as explanatory variables. The Firth penalized likelihood method will be used to reduce bias in the parameter estimates and to avoid non-convergence due to quasi separation.

Observed values and derived change/percent change from baseline values in HAM-D subscale and individual item scores at each time point will be summarized by treatment group. The change from baseline in HAM-D subscale and individual item scores will be analyzed by MMRM methods similar to the primary endpoint, as described in Section 9.3.1.3.

Bar charts showing the percentage of subjects with HAM-D response and remission by treatment and visit will be provided.

9.3.2.2 Montgomery-Åsberg Depression Rating Scale (MADRS)

Description of MADRS

The MADRS is a 10-item diagnostic questionnaire that psychiatrists use to measure the severity of depressive episodes in subjects with mood disorders. It was designed as an adjunct to the HAM-D that would be more sensitive to the changes brought on by antidepressants and other forms of treatment than the Hamilton Scale.

The questionnaire includes questions on the following symptoms: apparent sadness; reported sadness; inner tension; reduced sleep; reduced appetite; concentration difficulties; lassitude; inability to feel; pessimistic thoughts; and suicidal thoughts. If more than 2 individual items are missing, the MADRS total score will not be calculated and will be left as missing. If less than or equal to 2 individual item scores are missing, the missing item scores will be imputed by the mean of all other available item scores. Imputed individual scores will be rounded to the nearest integer.

Higher MADRS scores indicate more severe depression, and each item yields a score of 0 to 6. The MADRS total score will be calculated as the sum of the 10 individual item scores. The overall score ranges from 0 to 60 (McDowell 2006; Müller-Thomsen 2005).

The MADRS will be collected in the clinic at Screening; before dosing on Day 1; 12 hours ± 1 hour following the evening dose on Days 2 (before PA4 only), 3, 8, and 15; and in the morning on Days 21 and 45.

Analysis of MADRS

Observed values and derived change/percent change from baseline values in MADRS total score and individual item scores at each time point will be summarized by treatment group. The change from baseline in MADRS total score will be analyzed by MMRM methods similar to the primary endpoint, as described in Section 9.3.1.3. The LS mean change from baseline (and 95% CI of the mean) in MADRS total score based on MMRM results will be plotted over time by treatment group.

The MMRM analysis for change from baseline in MADRS total score analysis will be repeated for the following subgroups: subjects on antidepressant medications at treatment start and subjects who are not on antidepressant medications at treatment start.

Description of MADRS Response and Remission

The MADRS response is defined as a 50% or greater reduction from baseline in MADRS total score (ie, percent change from baseline ≤-50%); MADRS remission is defined as a MADRS total score of ≤10.

Analysis of MADRS Response and Remission

The MADRS response and remission results will be analyzed by GEE and logistic regression methods similar the HAM-D response and remission, as described in Section 9.3.2.1.

9.3.2.3 Clinical Global Impression (CGI)

Description of CGI

The CGI is a validated measure often utilized in clinical trials to allow clinicians to integrate several sources of information into a single rating of the subject's condition. The CGI scale consists of 3 items. Only the first 2 items are being used in this study. Only the first item of the CGI, the CGI-S item, will be administered before treatment.

The CGI-S uses a 7-point Likert scale to rate the severity of the subject's illness at the time of assessment, relative to the clinician's past experience with subjects who have the same diagnosis. Considering total clinical experience, a subject is assessed on severity of mental illness at the time of rating as 1=normal, not at all ill; 2=borderline mentally ill; 3=mildly ill; 4=moderately ill; 5=markedly ill; 6=severely ill; and 7=extremely ill.

The CGI-I employs a 7-point Likert scale to measure the overall improvement in the subject's condition post-treatment. The Investigator will rate the subject's total improvement whether or not it is due entirely to drug treatment. Response choices include: 0=not assessed, 1=very much improved, 2=much improved, 3=minimally improved, 4=no change, 5=minimally worse, 6=much worse, and 7=very much worse. The CGI-I is only rated at post-treatment assessments. By definition, all CGI-I assessments are evaluated against baseline conditions. The CGI-I response will be defined as having a CGI-I score of "very much improved" (score of 1) or "much improved" (score of 2).

The CGI-S will be collected in the clinic at Screening; before dosing on Day 1; 12 hours ± 1 hour following the evening dose on Days 2 (before PA4 only), 3, 8, and 15; and in the morning on Days 21 and 45.

The CGI-I will be collected in the clinic 12 hours ± 1 hour following the evening dose on Days 2 (before PA4 only), 3, 8, and 15; and in the morning on Days 21 and 45.

Analysis of CGI

All CGI-S and CGI-I scores will be summarized as both categorical (ie, number and percentages) and continuous (ie, descriptive statistics) at each time point (Screening and Days 1, 3, 8, 15, 21, and 45) by treatment group. The mean CGI-I score will be plotted over time by treatment group.

The CGI-I response will be analyzed using GEE and logistic regression similar to HAM-D response and remission as described in Section 9.3.2.1. Baseline CGI-S score will be included in the model.

The GEE analysis for CGI-I response analysis will be repeated for the following subgroups: subjects on antidepressant medications at treatment start and subjects who are not on antidepressant medications at treatment start.

A bar chart showing the percentage of subjects with CGI-I response by treatment and visit will be provided.

9.3.2.4 Hamilton Anxiety Rating Scale (HAM-A)

Description of HAM-A

The 14-item HAM-A will be used to rate the severity of symptoms of anxiety (Hamilton 1959). Each of the 14 items is defined by a series of symptoms, and measures both psychic anxiety (mental agitation and psychological distress) and somatic anxiety (physical complaints related to anxiety). Scoring for HAM-A is calculated by assigning scores of 0 (not present) to 4 (very severe), with a total score (calculated as the sum of the 14 individual scores) ranging from 0 to 56, where ≤17 indicates mild severity, 18 to 24 indicates mild to moderate severity, 25 to 30 indicates moderate to severe severity, and >30 indicates severe severity. If more than 2 individual items are missing, the HAM-A total score will not be calculated and will be left as missing. If less than or equal to 2 individual item scores are missing, the missing item scores will be imputed by the mean of all other available item scores.

The HAM-A will be collected in the clinic before dosing on Day 1; 12 hours ± 1 hour following the previous evening's dose on Days 3, 8, and 15; and in the morning on Days 21 and 45.

Analysis of HAM-A

Observed values and change/percent change from baseline in HAM-A total score at each time point will be summarized by treatment group. The change from baseline in HAM-A total score will be analyzed by MMRM methods similar to the primary endpoint, as described in Section 9.3.1.3. Additionally, number and percentage of subjects in each HAM-A total score category at each time point will be presented by treatment group. The LS mean change from baseline (and 95% CI of the mean) in HAM-A total score based on MMRM results will be plotted over time by treatment group.

The MMRM analysis for change from baseline in HAM-A total score analysis will be repeated for the following subgroups: subjects on antidepressant medications at treatment start and subjects who are not on antidepressant medications at treatment start.

9.3.2.5 Beck Anxiety Inventory (BAI)

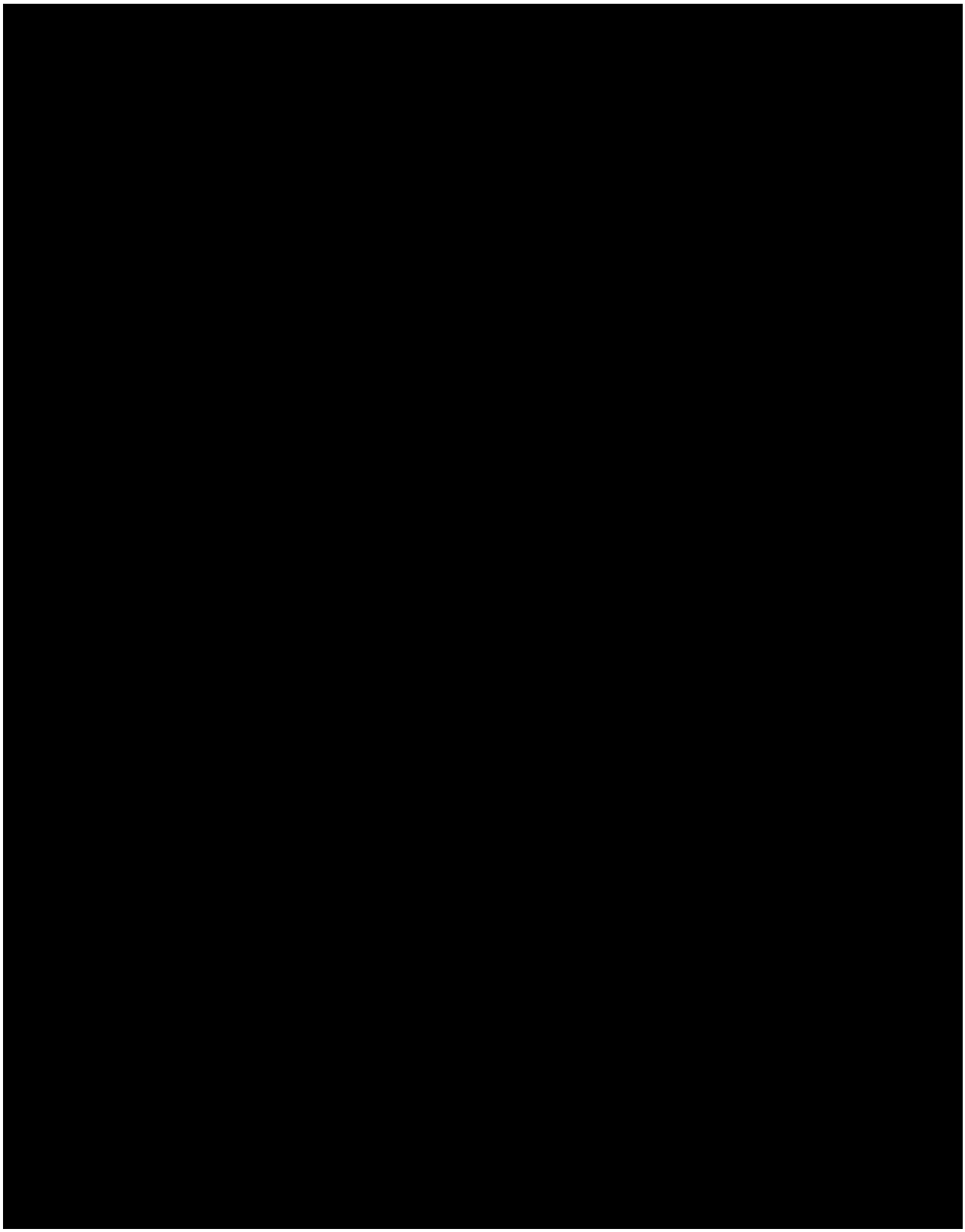
Description of BAI

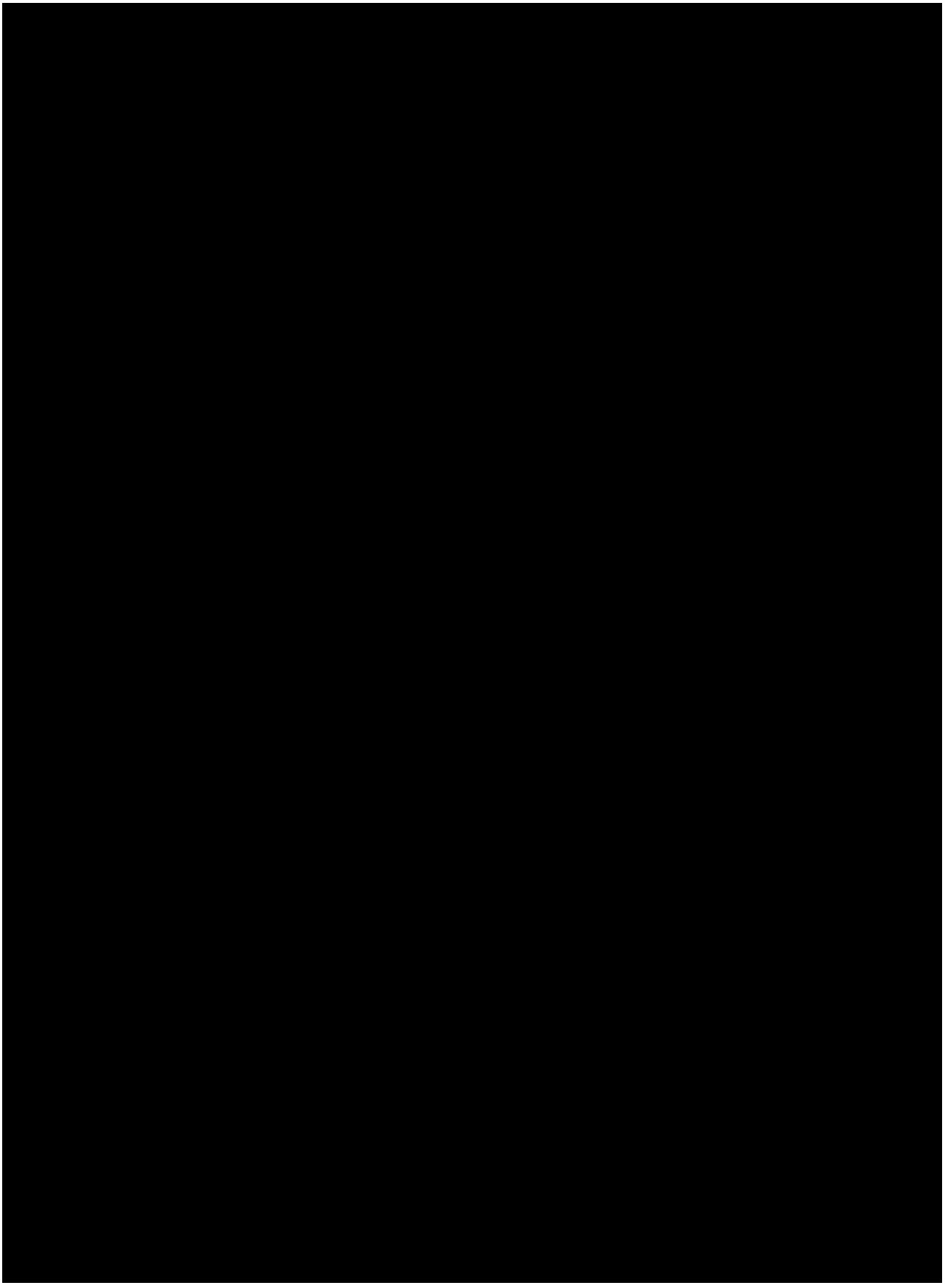
The BAI is a subject-rated measure of anxiety with 21 items. Scoring for BAI is calculated by assigning scores of 0, 1, 2, and 3 to the response categories of "not at all," "mildly," "moderately," and "severely", respectively. The BAI total score (calculated as the sum of the 21 individual item scores) ranging from 0 to 48, where a score of 0 to 21 indicates low anxiety, 22 to 35 indicates moderate anxiety, and 36 and above indicates potentially concerning levels of anxiety (Beck 1988).

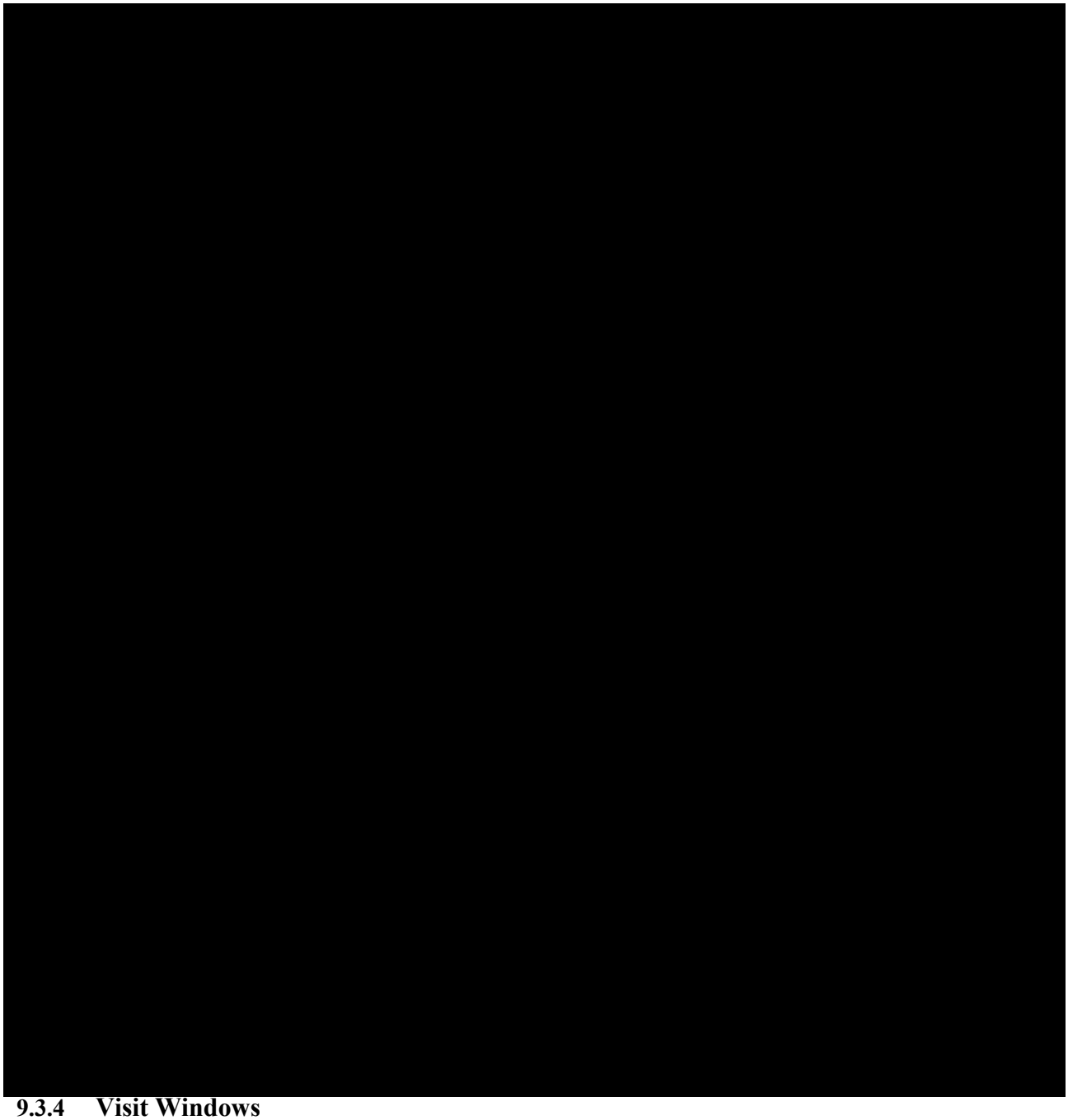
The BAI will be collected in the clinic at Screening; before dosing on Day 1; 12 hours ± 1 hour following the evening dose on Days 8 and 15; and in the morning on Days 21 and 45. All assessments are to be

completed within ±30 minutes of the scheduled time point. The BAI was only collected for subjects enrolled before PA 4.

Anxiety data collected on the BAI will be listed.







9.3.4 Visit Windows

Any unscheduled or early termination (ET) visit will be mapped to a scheduled visit for analysis using the date/time of collection/assessment as a basis to determine study day and time, which will be mapped to the intended visit. Once analysis visit windows are assigned, all visits (including scheduled visits, unscheduled visits, and ET visits) will be eligible for being flagged as the "analyzed record" within the analysis window. The analysis visit windows are as follows:

Analysis Visit Window Name	Target Day of Visit	Lower Bound of Window (day)	Upper Bound of Window (day)
Baseline	1	--	1
Day 3 (± 1)	3	2	4
Day 8 (± 1)	8	7	9
Day 15 (± 1)	15	14	16
Day 21 (± 1)	21	20	22
Day 45 (± 3)	45	42	48

A subject's individual analysis visit window could potentially contain more than 1 visit. In the event of multiple visits falling within an analysis window or in case of a tie, the following rules will be used in sequence to determine the "analyzed record" for the analysis visit window:

- If there is a scheduled visit for the analysis visit window, the scheduled visit data will be used.
- If there is no scheduled visit for the analysis visit window (eg, a subject missed the visit or the assessment was not collected), the visit closest to the scheduled study day will be used.
- If there is no scheduled visit for the analysis visit window and there are visits equally close to the day of the scheduled visit, the latest visit will be used.

All visits, including visit not flagged as the "analyzed record," will be included in the subject listings.

9.3.5 Pooling of Centers

For supporting primary efficacy analysis purposes, centers with fewer than 15 subjects per center will be pooled, first centers with fewer than 15 subjects will be ranked by size. Starting with the largest site, centers will be pooled until 15 or more subjects are reached to create a unique regional center. Continuing this process of creating unique regional centers until there are less than 15 subjects remaining. The remaining centers will then be added to the last unique regional center created.

Unless the combined number of subjects from all centers with fewer than 15 subjects in a region is fewer than 15 subjects, all pooled centers should have at least 15 subjects.

9.3.6 Protocol Deviations

Prior to database lock, major protocol deviations are defined as those deviations from the study protocol that may have the ability to impact the efficacy results. As such, subjects with one or more major deviation will be excluded from the per-protocol set (as defined in Section 8.3).
Major protocol deviations include (but are not limited to):

- Deviation from inclusion/exclusion criteria
- Withdrawal criteria met during the study but subject was not withdrawn
- Prohibited concomitant medications
- Substantial deviations from the dosing schedule
- Had overall treatment compliance rate < 80% or > 120%
- Other

Where possible, major deviations will be programmatically verified. Listings of concomitant medications in Excel format will be reviewed and to be identified for prohibited medications prior to unblinding. In

addition, listings of protocol deviations will be provided to the sponsor for review and classification of each protocol deviation as major/minor. These classifications will then be merged with the protocol deviations data for inclusion in tables and listings.

9.4 Safety Analysis

Safety will be evaluated through frequency and severity of AEs, as well as observed values and changes from baseline in vital signs, clinical laboratory measures, ECG parameters, suicidal ideation using the C-SSRS, and SSS.

All safety data will be presented in individual subject data listings. Safety analysis described below will be conducted using the Safety Set and treatment received (SAGE-217 if any dose of SAGE-217 is received, or placebo if all doses received are placebo) as the treatment group.

Safety Evaluation	Incidence	Observed Value	Change from Baseline	Abnormality/ Clinical Significance	Potentially Clinical Significance (PCS)
AEs	X				
CMs [1]	X				
Labs		X	X	*	x
ECG		X	X	*	x
Vital Signs		X	X		x
C-SSRS		X	X		

X = Safety Assessment will be summarized in tables

* = Safety Assessment will be presented in individual subject data listings; abnormal clinically significant lab, ECG, and vital findings will only be noted on AE CRFs

[1] Including antidepressant medication usage

9.4.1 Adverse Events

Adverse events (AEs) will be coded using the MedDRA coding system (version 19.1 or higher). The analysis of AEs will be based on the concept of treatment emergent AEs (TEAEs). A TEAE is defined as an AE with onset after the start of study drug, or any worsening of a pre-existing medical condition/AE with onset after the start of study drug.

A summary of TEAEs will be provided. Frequencies and percentages of the following will be included:

- Any TEAE
- Severe TEAEs
- TEAEs which are possibly or probably related to study drug
- TEAEs Leading to Study Drug Discontinuation
- TEAEs Leading to Study Discontinuation

- SAEs
- AEs Leading to Death

The incidence of all TEAEs, TEAEs by maximum severity, TEAEs by greatest relationship (related>non-related, related = probably or possibly related or missing) to study drug, and all SAEs (if ≥10 subjects experience an SAE) will be summarized by MedDRA System Organ Class (SOC), PT, and treatment received. Additionally, the incidence of all TEAEs will be summarized by PT and treatment received. Incidences will be presented in order of decreasing frequency and then alphabetically, first SAGE-217 then placebo.

All AEs, SAEs, AEs leading to study drug discontinuation, and AEs leading to death through Day 45 follow-up visit will be presented as separate listings.

9.4.2 Clinical Laboratory

All summaries of laboratory values will be presented using SI units. Observed hematology, chemistry, and urinalysis values and changes from baseline at each time point will be summarized by treatment received. Shifts from baseline regarding out-of-normal range values will also be summarized by treatment received. The shift from baseline status of low/normal/high to low or high at any time during treatment, any time during the study and at the end of the study will be presented. All clinical laboratory results will be listed by subject and timing of collection. This listing will include data from scheduled and unscheduled time points. Clinically significant abnormal findings will be reported as AEs in the clinical database, hence will be reported as part of AE reporting. Pregnancy test results will be listed only.

Potentially clinically significant (PCS) values for selected hematology and clinical chemistry parameters have been identified in the following table. The number and percentage of subjects with PCS values at each post-baseline visit, any time during treatment, any time during the study, and at the end of the study will be summarized by treatment received.

Hematology Potentially Clinically Significant Values

Hematology parameter	Potentially clinically significant values
Hemoglobin	<125 g/L or >185 g/L (males) <110 g/L or >165 g/L (females)
Hematocrit (HCT)	<0.415 or >0.504 (males) <0.359 or >0.446 (females)
Platelet count	<125 10^9/L or >600 10^9/L
Leukocytes	<2.5 10^9/L or >15 10^9/L
Basophils	>0.5 10^9/L
Eosinophils	>1.5 10^9/L
Lympocytes	<0.5 10^9/L or >6.0 10^9/L
Monocytes	>1.4 10^9/L
Neutrophils	<1.5 10^9/L

Chemistry Potentially Clinically Significant Values

Chemistry parameter	Potentially clinically significant values
Alanine aminotransferase (ALT)	>3 x ULN
Albumin	<28 g/L or >70 g/L
Aspartate aminotransferase (AST)	>3 x ULN
Alkaline Phosphatase (ALP)	>1.5 x ULN
Bilirubin	>2 x ULN
Blood urea nitrogen (BUN)	>10.71 mmol/L

Calcium	<2.0 mmol/L or >2.75 mmol/L
Chloride	<90 mmol/L or >120 mmol/L
Creatinine	>140 umol/L
Phosphate	<2 mg/dL or >5 mg/dL
Potassium	<3.5 mmol/L or >5.2 mmol/L
Sodium	<132 mmol/L or >145 mmol/L
Protein	<45 g/L
Glucose	<2.8 mmol/L or >13.9 mmol/L

Summaries of abnormal liver enzymes and liver function tests post-baseline will be presented for ALT, AST, Total Bilirubin, and ALP. Shifts in ALT and AST baseline values to maximum post-baseline values will be summarized using the categories ≤1x ULN, >1x to 3x ULN, >3x to 5x ULN, >5x to 10x ULN, and >10x ULN.

All scheduled and unscheduled results will be included in the laboratory PCS analyses.

9.4.3 Vital Signs

Vital sign results (body temperature, heart rate, respiratory rate, supine and standing diastolic blood pressure, and supine and standing systolic blood pressure) will be listed by subject and timing of collection. Change from baseline for diastolic blood pressure and systolic blood pressure will be calculated separately for supine and standing position. Observed values and change from baseline at each time point will be summarized by treatment received.

Additionally, the change in blood pressure from supine to standing position (standing – supine) will be summarized. A listing will be provided showing the difference between supine and standing blood pressures.

Post-baseline vital signs meeting any of the criteria below for low, high or change from baseline will be identified as potentially clinically significant (PCS). These will be summarized and listed by subject:

Potentially Clinically Significant Vital Sign Values

Parameter	Low	High	Change from Baseline
Systolic Blood Pressure (mmHg)	<90	>180	+ or – change of ≥30
Diastolic Blood Pressure (mmHg)	<50	>110	+ or – change of ≥20
Heart Rate (beats/min)	<40	>120	Not defined

In addition, the number of subjects with post-baseline orthostatic hypotension, defined as either a decrease in systolic blood pressure of ≥20 mmHg or a decrease in diastolic blood pressure of ≥10 mmHg from supine to standing position, will be summarized.

9.4.4 Electrocardiogram

The following 12-lead ECG parameters will be listed for each subject: heart rate, PR interval, QRS duration, QT interval, and QT interval using Fridericia's correction for heart rate (QTcF, derived as QT interval / (RR interval)$^{1/3}$). Any clinically significant abnormalities or changes in ECGs should be recorded as an AE. Observed values and change from baseline at each time point will be summarized by treatment received. Additionally, the number and percentage of subjects with a QTcF value meeting following clinical significance criteria will be summarized separately by treatment received.

- Maximum value >450 to 480 msec
- Maximum value >480 to 500 msec

- Maximum value >500 msec
- Maximum increase from baseline >30 to 60 msec
- Maximum increase from baseline >60 msec

9.4.5 Physical Examination

Screening physical examination results that are clinically significant will be collected in the medical history eCRF page, and hence will be included in medical history listing. Post-screening physical examination results that are clinically significant will be reported as AEs in the clinical database, hence will be reported as part of AE reporting. A listing will be provided containing information on when physical examination occurred and reason it was not done if applicable.

9.4.6 Columbia-Suicide Severity Rating Scale (C-SSRS)

Suicidality will be monitored during the study using the C-SSRS (Posner 2011). This scale consists of a baseline evaluation that assesses the lifetime experience of the subject with suicidal ideation and behavior, and a post-baseline evaluation that focuses on suicidality since the last study visit. The C-SSRS includes "yes" or "no" responses for assessment of suicidal ideation and behavior as well as numeric ratings for severity of ideation, if present (from 1 to 5, with 5 being the most severe). Assign a score of 0 if no ideation is present.

Suicidal Ideation items include:

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 items include:

Category 6 – Preparatory Acts or Behavior
Category 7 – Aborted Attempt
Category 8 – Interrupted Attempt
Category 9 – Actual Attempt (non-fatal)
Category 10 – Completed Suicide
Self-injurious behavior without suicidal intent is also a C-SSRS outcome (although not suicide-related) and has a binary response (yes/no).

Endpoints:

Composite endpoints based on the above categories are defined below:

- Suicidal **ideation**: A "yes" answer at any time during treatment to any one of the five suicidal ideation questions (Categories 1-5) on the C-SSRS.
- Suicidal **behavior**: A "yes" answer at any time during treatment to any one of the five suicidal behavior questions (Categories 6-10) on the C-SSRS.
- Suicidal **ideation or behavior**: A "yes" answer at any time during treatment to any one of the ten suicidal ideation and behavior questions (Categories 1-10) on the C-SSRS.

Baseline is defined as the most severe ideation and behavior reported in the past 24 months prior to the first dose of study drug. Suicidality data collected on the C-SSRS will be listed for all subjects. Tables will include results from the Suicidal Ideation and Suicidal Behavior sections of the C-SSRS. Frequencies and percentages of subjects with a response of "Yes" at any point as well as by study visit on the Suicidal Ideation and Suicidal Behavior items will be summarized by treatment received.

Additionally, shifts of whether subjects experienced suicidal ideation, behavior, or both from pre-treatment to post-treatment will be summarized in a shift table. An additional shift table of maximum ideation from pre-treatment to post-treatment will also be provided.

A bar-chart summarizing the number of subjects with suicidal ideation, suicide attempts, and suicide by visit will be provided.

9.4.7 Stanford Sleepiness Scale (SSS)

The SSS is a subject-rated scale designed to quickly assess how alert a subject is feeling. Degrees of sleepiness and alertness are rated on a scale of 1 to 7, where the lowest score of 1 indicates the subject is "feeling active, vital, alert, or wide awake" and the highest score of 7 indicates the subject is "no longer fighting sleep, sleep onset soon; having dream-like thoughts". The SSS was only collected for subjects enrolled before PA4.

Sedation data collected on the SSS will be listed for all applicable subjects.

[redacted]

10 SUMMARY OF INTERIM ANALYSES

Not applicable.

11 REFERENCES

Barkin, J. L. (2009, June 12). The Development and Testing of the Barkin Index of Maternal Functioning. Pittsburgh, PA.

Beck, A., Epstein, N., Brown, G., & Steer, R. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *J Consult Clin Psychol*, 56:893-7.

Cox, J., Holden, J., & Sagovsky, R. (1987). Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. *Br J Psychiatry*, 150:782-6.

Hamilton, M. (1959). The assessment of anxiety states by rating. *Br J Med Psychol*, 32:50-5.

Hamilton, M. (1960). A rating scale for depression. *J Neurol Neurosurg Psychiatry*, 23:56-62.

Kroenke, K., Spitzer, R., & Williams, J. (2001). The PHQ-9. *J Gen Intern Med*, 16(9):606-613.

McDowell, I. (2006). *Measuring Health: A guide to rating scales and questionnaires.* New York: Oxford University Press.

Müller-Thomsen, T., Arlt, S., Mann, U., & al. (2005). Detecting depression in Alzheimer's disease: evaluation of four different scales. *Archives of Clinical Neuropsychology*, 20:271-6.

Posner, K., Brown, G., Stanley, B., & al. (2011). The Columbia-Suicide Severity Rating Scale: initial validity and consistency findings from three multisite studies with adolescents and adults. *Am J Psychiatry*, 168(12):1266-77.

[redacted]

12 LIST OF APPENDICES

12.1 Appendix A: Schedule of Events (subjects enrolled before PA4)

	Screening Period	**Treatment Period**														**Follow-up Period**		
Visits	**SCREENING**	**OUTPATIENT**														**OUTPATIENT**		
Visit Days	**D-14 to -1**	**D1***	**D2**	**D3**	**D4**	**D5**	**D6**	**D7**	**D8**	**D9**	**D10**	**D11**	**D12**	**D13**	**D14**	**D15**	**D21/ET (+1d)**	**D45 (+3d)**
Study Procedure																		
Informed Consent	X																	
Inclusion/Exclusion	X	X																
Demographics	X																	
Medical/Family History [a]	X																	
Physical Examination	X								X[b]							X**Error! Reference source not found.**	X	
Body Weight/Height	X															X[b,c]	X[c]	
Clinical Laboratory Assessments [d]	X								X[b]							X[b]	X	
Drug & Alcohol Screen [e]	X	X																
Pregnancy Test [f]	X	X															X (ET only)	X
Hepatitis & HIV Screen	X																	
Hormone Sample [g]	O								O[b]							O[b]		
Genetic Sample [h]	O																	
Vital Signs [i]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X[b]	X	X
Pulse Oximetry [j]		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

	Screening Period	Treatment Period														Follow-up Period		
Visits	**SCREENING**	**OUTPATIENT**														**OUTPATIENT**		
Visit Days	**D-14 to -1**	**D1***	**D2**	**D3**	**D4**	**D5**	**D6**	**D7**	**D8**	**D9**	**D10**	**D11**	**D12**	**D13**	**D14**	**D15**	**D21/ET (+1d)**	**D45 (+3d)**
Study Procedure																		
12-Lead ECG [k]	X	X	X						X							X	X	
C-SSRS [l]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X[b]	X	X
CGI-I [m]			X	X					X							X	X[b]	X[b]
CGI-S **Error! Reference source not found.**	X	X	X	X					X							X	X[b]	X[b]
HAM-A **Error! Reference source not found.**		X		X					X							X	X[b]	X[b]
HAM-D **Error! Reference source not found.**	X	X	X	X	X	X	X		X							X	X[b]	X[b]
MADRS **Error! Reference source not found.**	X	X	X	X					X							X	X[b]	X[b]
[redacted]																		
[redacted]																		
BAI **Error! Reference source not found.**	X	X							X							X	X[b]	X[b]
[redacted]																		
[redacted]																		
SSS [n]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X[b]		
[redacted]																		

	Screening Period	Treatment Period														Follow-up Period		
Visits	**SCREENING**	**OUTPATIENT**														**OUTPATIENT**		
Visit Days	**D-14 to -1**	**D1***	**D2**	**D3**	**D4**	**D5**	**D6**	**D7**	**D8**	**D9**	**D10**	**D11**	**D12**	**D13**	**D14**	**D15**	**D21/ET (+1d)**	**D45 (+3d)**
Study Procedure																		
[redacted]																		
Study Drug Administration [r]		X	X	X	X	X	X	X	X	X	X	X	X	X	X			
Adverse Events	X																	
Prior/Concomitant Medications [s]	X																	

BAI = Beck Anxiety Inventory; [redacted]; CGI-I = Clinical Global Impression of Improvement; CGI-S – Clinical Global Impression - Severity; C-SSRS = Columbia Suicide Severity Rating Scale; ECG = electrocardiogram; [redacted]; ET = early termination; HAMA = Hamilton Anxiety Rating Scale; HAM-D = Hamilton Rating Scale for Depression, 17-item; [redacted]; HIV = human immunodeficiency virus; MADRS = -Montgomery-Åsberg Depression Rating Scale; O = Optional; [redacted]; SSS = Stanford Sleepiness Scale

*All D1 procedures are to be completed prior to dosing

Error! Reference source not found. Subjects will be specifically asked about the depression and anxiety diagnoses.

Error! Reference source not found. To be performed/collected in the morning at the clinic.

Error! Reference source not found. Weight only.

Error! Reference source not found. Safety laboratory tests will include hematology, serum chemistry, coagulation, select hormone parameters, and urinalysis. Laboratory assessments are to be completed at the Screening visit, in the morning on Day 8 and Day 15, and during the follow-up visit on Day 21.

Error! Reference source not found. Urine toxicology for selected drugs of abuse and serum or breath test for alcohol.

Error! Reference source not found. Serum pregnancy test at screening and urine pregnancy test at Day 1 and Day 45. A urine pregnancy test will also be collected as part of the early termination assessments for subjects who discontinue the study early.

Error! Reference source not found. An optional blood sample for stress hormone levels, kynurenine biochemistry, and markers of inflammation, where consent is given.

Error! Reference source not found. An optional genetic sample for biomarker testing, where consent is given.

Error! Reference source not found. Vital signs, including oral temperature (°C), respiratory rate, heart rate, and blood pressure (supine and standing), will be obtained within ±5 minutes of the scheduled time point through 0.5 hours after dosing, and ±10 minutes of the scheduled time point from 1 hour after dosing and greater, unless the subject is asleep between the hours of 11:00 PM and 06:00 AM. Vital signs will be completed at the following time points: predose (within 1 hour of dosing), 0.25, 0.5, and 1 hour after the dosing. Vital signs may be repeated at the discretion of the Investigator or Visiting Nurse as clinically indicated.

Error! Reference source not found. Pulse oximetry will be obtained within ±5 minutes of the scheduled time point through 0.5 hours after dosing, and ±10 minutes of the scheduled time point from 1 hour after dosing and greater, unless the subject is asleep between the hours of 11:00 PM and 06:00 AM. Pulse oximetry will be completed at the following time points: predose (within 1 hour of dosing), 0.25, 0.5, and 1 hour after the dosing. Pulse oximetry may be repeated at the discretion of the Investigator or Visiting Nurse as clinically indicated.

Error! Reference source not found. Performed at screening and within ±60 minutes of the scheduled time point at 9:00 AM on Days 1, 2, 8, and 15, and during the follow-up visit on Day 21.

Error! Reference source not found. The "Baseline/Screening" C-SSRS form will be completed at screening. The "Since Last Visit" C-SSRS form will be completed at all subsequent time points.

Error! Reference source not found. To be completed 12 hours ± 1 hour following the evening dose in the clinic at the scheduled time points.

Error! Reference source not found. SSS will be completed at the following time points: predose (within 1 hour of dosing), 0.25, 0.5, 1, and 12 hours after dosing. To be completed within ±15 minutes of the scheduled time point (except for the 12-hour time point, which is to be completed within ±30 minutes).

Error! Reference source not found. HRCU will be administered at screening (Screening version). HRCU (post-screening log) will be administered from Day 1 to Day 45 to collect a cumulative assessment of HRCU over the 45 days.

[redacted]

Dispense Study DrugDosing will occur daily at 8:00 PM ±15 minutes by the nursing service or at the clinic. Subjects will be instructed to take study drug with food.

Error! Reference source not found. To include those taken prior to the first dose of study drug and throughout the study, as well as history of antidepressant medications and treatment.

12.2 Appendix B: Schedule of Events (subjects enrolled as of PA4)

	Study Period / Visit						
	Screening Period	Treatment Period [a]				Follow-up Period	
Study Procedure	**D-28 to -1**	**D1***	**D3 (+1d)**	**D8**	**D15**	**D21/ET (±1d)**	**D45 (±3d)**
Clinic Visit	X	X	X	X	X	X	X
Informed Consent	X						
Inclusion/Exclusion	X	X					
Demographics	X						
Medical/Family History [b]	X						
Physical Examination	X			X[c]	X[c]	X	
Body Weight/Height	X				X[c, d]	X[d]	
Clinical Laboratory Assessments [e]	X			X[c]	X[c]	X	
Drug & Alcohol Screen [f]	X	X					
Pregnancy Test[g]	X	X				X (ET only)	X
Hepatitis & HIV Screen	X						
Hormone Sample [h]	O			O[c]	O[c]		
Genetic Sample [i]	O						
Vital Signs [j]	X	X	X	X	X[c]	X	X
12-Lead ECG [k]	X	X		X	X	X	
C-SSRS [l]	X	X	X	X	X[c]	X	X
CGI-I			X	X	X	X[c]	X[c]
CGI-S	X	X	X	X	X	X[c]	X[c]
HAM-A		X	X	X	X	X[c]	X[c]
HAM-D [m]	X	X	X	X	X	X[c]	X[c]

Study Procedure	Study Period / Visit						
	Screening Period	Treatment Period [a]				Follow-up Period	
	D-28 to -1	D1*	D3 (+1d)	D8	D15	D21/ET (±1d)	D45 (±3d)
MADRS	X	X	X	X	X	X[c]	X[c]
[redacted]							
Dispense Study Drug [q]		X		X			
Study Drug Administration [r]		X (QD until Day 14 - inclusive)					
Treatment Compliance Call [s]		X (QD until Day 14 - inclusive)					
Adverse Events	X						
Prior/Concomitant Medications [t]	X						

[redacted]; CGI-I = Clinical Global Impression of Improvement; CGI-S – Clinical Global Impression - Severity; C-SSRS = Columbia Suicide Severity Rating Scale; ECG = electrocardiogram; [redacted]; ET = early termination; HAMA = Hamilton Anxiety Rating Scale; HAM-D = Hamilton Rating Scale for Depression, 17-item; [redacted]; HIV = human immunodeficiency virus; MADRS = Montgomery-Åsberg Depression Rating Scale; O = Optional; [redacted]; QD = once daily; [redacted]
*All D1 procedures are to be completed prior to dosing

[a] An unscheduled visit may be needed if a dose adjustment is deemed necessary by the Investigator at any time during the treatment period in order for any remaining current dose to be returned and for the adjusted dose to be dispensed.

[b] Subjects will be specifically asked about the depression and anxiety diagnoses listed in **Error! Reference source not found.** of the protocol.
[c] To be performed/collected in the morning at the clinic.
[d] Weight only.
[e] Safety laboratory tests will include hematology, serum chemistry, coagulation, select hormone parameters, and urinalysis.
[f] Urine toxicology for selected drugs of abuse and serum (screening only) or breath test for alcohol.
[g] Serum pregnancy test at screening and urine pregnancy test at Day 1 and Day 45. A urine pregnancy test will also be collected as part of the early termination assessments for subjects who discontinue the study early.
[h] An optional blood sample for stress hormone levels, kynurenine biochemistry, and markers of inflammation, where consent is given.

[i] An optional genetic sample for biomarker testing, where consent is given.
[j] Vital signs include respiratory rate, oral temperature, and supine (for at least 5 minutes prior to the measurement) and standing systolic and diastolic blood pressure and heart rate. Vital signs may be repeated at the discretion of the Investigator as clinically indicated.
[k] Performed at screening and within ±60 minutes of the scheduled time point at 9:00 AM.
[l] The "Baseline/Screening" C-SSRS form will be completed at screening. The "Since Last Visit" C-SSRS form will be completed at all subsequent time points.
[m] HAM-D to be completed at the scheduled clinic visits 12 hours (±1 hour) following the evening dose.

[q] Per investigator discretion, an additional visit to dispense study drug may occur. The only planned procedure for this visit will be study drug dispensation.
[r] Dosing will occur daily at 8:00 PM ±30 minutes with food. If the dose is not administered within <60 minutes before/after the scheduled dose, the subject will skip the dose and take the next scheduled dose on the following day (a dose occurring >30 minutes but <60 minutes after the scheduled time will be considered a protocol deviation).
[s] A treatment compliance call will be made within approximately 1 hour following the scheduled evening dose on Days 1-14.
[t] To include those taken prior to the first dose of study drug and throughout the study, as well as history of antidepressant medications and treatment listed in **Error! Reference source not found.** of the protocol.

12.3 Appendix C: Details of Statistical Methodology

Sample SAS code for MMRM:

```
PROC MIXED DATA = xxx;
  CLASS  trt avisit usubjid AntiDpUe;
  MODEL change = base trt avisit avisit*trt AntiDpUe / ddfm=kr;
  REPEATED avisit/subject = usubjid type= un;
* if type= un does not converge, use type= TOEP;
* if type=TOEP does not converge, use type= cs;
* if type=cs does not converge, use type= ar(1);
  LSMEANS avisit*trt / diff=all cl alpha=0.05;
RUN;
```

*Note: AntiDpUe represents baseline antidepressant medication use variable

Sample SAS code for GEE:

```
PROC GENMOD DATA = xxx;
  CLASS trt avisit usubjid AntiDpUe;
  MODEL resp = base trt avisit avisit*trt AntiDpUe/ DIST= bin link=logit;
  REPEATED subject=usubjid / type = un;
  LSMEANS avisit*trt / diff exp cl;
RUN;
```

Sample SAS code for Logistic Regression:

```
PROC LOGISTIC DATA = xxx;
  BY avisit;
   CLASS trt usubjid AntiDpUe /PARAM=GLM;
   MODEL resp (event='1') = base center trt AntiDpUe/firth; Note: for CGI-S, add cgisbase independent variable as well
   LSMEANS trtan / diff oddsratio cl exp;
RUN;
```

Sample SAS code for Stratified Wilcoxon Rank Sum Test:

```
PROC FREQ DATA = xxx;
   WHERE trtan in (1, 2);
   TABLES AntiDpUe *trtan*chg/cmh2 scores=modridit noprint;
RUN;
```