Official Title: **A Phase 3, Multicenter, Double-Blind, Randomized, Placebo-Controlled Study Evaluating the Efficacy of SAGE-217 in the Treatment of Adult Subjects With Major Depressive Disorder**

NCT Number: NCT03672175

Document Date: SAP Version 2.0: 12 November 2019

9. STATISTICAL ANALYSIS PLAN

The original statistical analysis plan (SAP) was modified once. SAP Version 2 (12 November 2019) is provided.

SAGE THERAPEUTICS INCORPORATED

Statistical Analysis Plan

Methods

Protocol Number 217-MDD-301

STUDY TITLE: A PHASE 3, MULTICENTER, DOUBLE-BLIND, RANDOMIZED, PLACEBO-CONTROLLED STUDY EVALUATING THE EFFICACY OF SAGE-217 IN THE TREATMENT OF ADULT SUBJECTS WITH MAJOR DEPRESSIVE DISORDER

Author of SAP: [redacted]

Version: Version 2.0

Version Date of SAP: 12 November 2019

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

Authorization Signature Page

A Phase 3, Multicenter, Double-Blind, Randomized, Placebo-Controlled Study Evaluating the Efficacy of SAGE-217 in the Treatment of Adult Subjects with Major Depressive Disorder

Author:

[redacted]

Name: [redacted] MS

Date

Position: [redacted]
Company: [redacted]

Approved by:

[redacted]

PhD, [redacted]
Sage Therapeutics, Inc.

Date

[redacted]

MD, MBA, [redacted]
Sage Therapeutics, Inc.

Date

[redacted]

MPH, [redacted]
Sage Therapeutics, Inc.

Date

[redacted]

MD, [redacted]
Sage Therapeutics, Inc.

Date

1 TABLE OF CONTENTS

2 LIST OF ABBREVIATIONS

Abbreviation or specialist term	Explanation
AE	adverse event
AR	Autoregressive
ATC	anatomical therapeutic chemical
BMI	body mass index
BYOD	bring-your-own-device
CGI-I	Clinical Global Impression scale for improvement
CGI-S	Clinical Global Impression scale for severity
[redacted]	[redacted]
CS	Clinical significant
C-SSRS	Columbia Suicide Severity Rating Scale
DSM-5	Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
ECG	Electrocardiogram
eCRF	electronic case report form
EOT	end of treatment
ET	early termination
FAS	Full Analysis Set
[redacted]	[redacted]
FSH	follicle stimulating hormone
GEE	generalized estimating equation
HAM-A	Hamilton Rating Scale for Anxiety
HAM-D	Hamilton Depression Rating Scale
hCG	human chorionic gonadotropin
HCV	hepatitis C virus
HIV	human immunodeficiency virus
ICF	informed consent form
IRT	interactive response technology
ISI	Insomnia Severity Index
LFT	Liver Function Tests
MADRS	Montgomery-Åsberg Depression Rating Scale
MCS	mental component summary
MDD	major depressive disorder
mFAS	Modified Full Analysis Set
MGH ATRQ	Massachusetts General Hospital Antidepressant Treatment Response Questionnaire
MMRM	mixed effects model for repeated measures
MedDRA	Medical Dictionary for Regulatory Activities
PCS	Potentially clinically significant
PCSC	potentially clinically significant change
PHQ-9	Patient health Questionnaire
[redacted]	[redacted]
PRO	patient-reported outcome
PT	preferred term
PWC-20	20-item Physician Withdrawal Checklist

Abbreviation or specialist term	Explanation
QTcF	QT corrected according to Fridericia's formula
SAE	serious adverse event
SAP	statistical analysis plan
SCID-5-CT	Structured Clinical Interview for Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition for clinical trials
SD	standard deviation
SE	standard error
SF-36	36-item Short Form survey
SI	International System of Units
sNAW	subjective number of awakenings
SOC	System Organ Class
sSL	subjective sleep latency
sTST	subjective total sleep time
sWASO	subjective wake after sleep onset
TEAE	treatment-emergent adverse event
UN	Unstructured
WHO-DD	World Health Organization – Drug Dictionary

3 INTRODUCTION

This statistical analysis plan (SAP) is for the final analysis of 217-MDD-301 study, and is based on clinical study protocol, version 5.0, dated 25 March 2019.

The purpose of the SAP is to describe in detail the statistical methodology and the statistical analyses to be conducted for the above-mentioned protocol. The SAP will be approved and finalized before database lock.

The first database lock and treatment unblinding will happen when all subjects with at least one dose of study drug complete the treatment period and follow up visit until Day 42 (or 4 weeks after the last dose). This will be called the Day-42 Database Lock; it is understood that some subjects will complete the study by this lock, some will be ongoing in the follow up period, hence the data beyond Day 42 will be incomplete. Second and final database lock is planned when all subjects with at least one dose of study drug complete the study (including the last follow up visit – planned for Day 42 for subjects in earlier protocol versions, and Day 182 for subjects for later protocol versions). This will be called the End of Study Database Lock.

4 STUDY OBJECTIVES

4.1 Primary Objective

The primary objective of Study 217-MDD-301 is to evaluate the efficacy of SAGE-217 in the treatment of major depressive disorder (MDD) compared to placebo.

4.2 Secondary Objective

The secondary objectives of Study 217-MDD-301 are:

- To evaluate the effect of SAGE-217 on sleep.
- To assess patient-reported outcome (PRO) measures as they relate to health-related quality of life and depressive symptoms.

4.3 Safety Objective

- To evaluate the safety and tolerability of SAGE-217.

[redacted]

5 STUDY ENDPOINTS

5.1 Efficacy Endpoints

- The primary efficacy endpoint is the change from baseline in the 17-item Hamilton Rating Scale for Depression (HAM-D) total score at Day 15. The estimand is the mean change.
- The secondary efficacy endpoints:
    - Change from baseline in the 17-item HAM-D total score at all other time points;

- HAM-D response, defined as a ≥50% reduction in HAM-D score from baseline, at Day 15 and all other time points;
- HAM-D remission, defined as HAM-D total score ≤7, at Day 15 and all other time points;
- Clinical Global Impression - Improvement (CGI-I) response, defined as "much improved" or "very much improved", at Day 15 and all other time points;
- Change from baseline in Clinical Global Impression - Severity (CGI-S) score at Day 15 and all other time points;
- Change from baseline in Hamilton Anxiety Rating Scale (HAM-A) total score at Day 15 and all other time points;
- Change from baseline in the Montgomery-Åsberg Depression Rating Scale (MADRS) total score at Day 15 and all other time points;
- Change from baseline in HAM-D subscale and individual item scores at all time points;
- Change in sleep at Day 15 and all other time points, as assessed by:
  - Insomnia Severity Index (ISI)
  - Subjective sleep parameters collected with the Core Consensus Sleep Diary
- Change from baseline in patient-reported outcome measures of health-related quality of life, as assessed by responses to the 36-item Short Form survey (SF-36) version 2, and of depressive symptoms, as assessed by the 9-item Patient Health Questionnaire (PHQ-9)

5.2 Safety Endpoints

- Incidence and severity of adverse events/serious adverse events;
- Changes from baseline in clinical laboratory measures, vital signs, and electrocardiogram (ECGs);
- Suicidal ideation and behavior using the Columbia Suicide Severity Rating Scale (C-SSRS);
- Potential withdrawal symptoms using the 20-item Physician Withdrawal Checklist (PWC-20)

6 STUDY DESIGN

6.1 Overall Design

Study 217-MDD-301 is a randomized, double-blind, parallel-group, placebo-controlled study in subjects with major depressive disorder (MDD). Randomization will be stratified based on use of antidepressant treatment (current/stable or not treated/withdrawn ≥60 days) at baseline and carried out

within each stratum in a 1:1:1 ratio to receive SAGE-217 20 mg, SAGE-217 30 mg, or matching placebo; subjects will be treated for 14 days beginning on Day 1.

The study will consist of a Screening Period of up to 28 days, a 14-day Treatment Period, and up to 6 months (168 days) of Follow-up.

The Screening Period begins with the signing of the informed consent form (ICF) at the Screening Visit; the ICF must be signed prior to beginning any screening activities. The diagnosis of MDD must be made according to Structured Clinical Interview for Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) Clinical Trial Version (SCID-5-CT) performed by a qualified healthcare professional. Subjects will undergo preliminary screening procedures at the Screening Visit to determine eligibility, including completion of the MADRS, HAM-D, and CGI-S.

Antidepressants are permitted provided subjects are on a stable dose for at least 60 days prior to Day 1 and agree to continue on the stable dose through the follow-up period (Day 42). Initiation of new antidepressants or any other medications that may potentially have an impact on efficacy or safety endpoints will not be allowed between screening and completion of the Day 42 assessments.

Eligible subjects will be stratified based on use of antidepressant treatment (current/stable or not treated/withdrawn ≥60 days) and randomized within each stratum to one of 3 treatment groups (SAGE-217 20 mg, SAGE-217 30 mg, or matching placebo) in a 1:1:1 ratio. Subjects will self-administer a single dose of study drug once daily in the evening with food, on an outpatient basis, for 14 days. Subjects will return to the study center during the treatment and follow-up periods as outlined Appendix A.

Dose reductions are not permitted. Subjects who cannot tolerate study drug will be discontinued from study drug and will receive treatment as clinically indicated. Subjects who discontinue treatment early should return to the site for an end of treatment (EOT) visit as soon as possible, preferably the day after treatment is discontinued. Follow-up visits should take place every 7 days after the last dose of treatment for 4 follow-up visits. If at any time after the EOT visit, a subject decides to terminate the study, the subject should return for an early termination (ET) visit. The EOT and ET visits can be on the same day if a subject discontinues study drug and terminates the study on the same day during a clinic visit; in this case, all assessments scheduled for both visits will be conducted, without repetition of any assessment

6.2 Sample Size and Power

Assuming a two-sided alpha level of 0.05, a sample size of 399 evaluable subjects 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 10 points. Assuming an 11% dropout rate and a 1:1:1 randomization ratio within each stratum (antidepressant use at baseline, yes or no), approximately 450 total randomized subjects will be required to obtain 399 evaluable subjects. Evaluable subjects are defined as those randomized subjects who receive study drug and have valid baseline and at least 1 post-baseline HAM-D assessment. Additional subjects may be randomized if the dropout rate is greater than 11%.

6.3 Randomization

This is a randomized, double-blind, placebo-controlled study. Subjects who meet the eligibility criteria will be randomized in a stratified manner based on use of antidepressant treatment (current/stable or not treated/withdrawn ≥60 days) at baseline; randomization will be done within each stratum in a 1:1:1 ratio to receive SAGE-217 20 mg, SAGE-217 30 mg, or matched placebo.

Randomization schedules will be generated by an independent statistician. The allocation to treatment group (SAGE-217 20 mg, SAGE-217 30 mg, or placebo) will be based on the randomization schedule. Randomization will be performed centrally via an interactive response technology (IRT) system.

6.4 Blinding and Unblinding

This is a randomized, double-blind, placebo-controlled study. Subjects, clinicians, and the study team will be blinded to treatment allocation during the study. The randomization schedules will be kept strictly confidential, accessible only to authorized personnel until the time of unblinding.

In case of medical emergency, the Investigator may request unblinding of an individual subject's treatment in the study via the IRT. 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 has been unblinded, the subject will be permanently discontinued from the study.

7 MODIFICATIONS

7.1 Modifications to the Approved Clinical Study Protocol

[redacted]

The original protocol and the subsequent 3 amendments had inclusion criterion of MADRS >= 30 and no criterion for HAM-D total score. A blinded review of data suggested that a substantial number of subjects had entered the study with Day 1 HAM-D scores <22, with some as low as 13, which did not correlate with the intended severity of depression to be treated in this study. Therefore, in amendment 4 the inclusion criterion was changed to: MADRS score >=32 AND HAM-D total score >=22. To make up for the subjects with low HAM-D total score at Day 1, the study may overenroll beyond 450 subjects.

Modified Full Analysis Set has been defined in Section 8.4 to correspond to targeted patient population; this set will be used in efficacy analysis.

Further clarification has been provided about the estimand in the primary efficacy analysis, including the target population, the variable of interest, the population level summary and identifying and dealing of intercurrent events.

7.2 Modifications to the Approved Statistical Analysis Plan

This is the first version of the SAP for the final analysis.

7.3 Modifications to the Approved DMC Charter

Not applicable.

8 ANALYSIS SETS

8.1 Randomized Set

The Randomized Set is defined as all subjects who are randomized.

8.2 Safety Set

The Safety Set is defined as all subjects who are administered study drug.

8.3 Full Analysis Set

The Full Analysis Set (FAS) is defined as all randomized subjects with valid baseline HAM-D and at least 1 post-baseline HAM-D evaluation.

8.4 Modified Full Analysis Set

The Modified Full Analysis Set (mFAS) is defined as all subjects in the FAS who have had a total HAM-D score of at least 22 at baseline.

[redacted]

8.6 Per Protocol Set

The Per Protocol Set is defined as all subjects in the FAS without any major protocol deviations related to efficacy. For further details, see Section 9.2.2.

9 STATISTICAL ANALYSIS

9.1 General Considerations

Unless otherwise specified, continuous endpoints will be summarized with n, mean standard deviation (SD), median, minimum (min) and maximum (max). If the measurements in the source (raw) data are integers, then the corresponding mean and median will be presented to 1 decimal place and the SD to 2 decimal places; if the measurements are obtained to 1 decimal place, then the mean and median will be presented to 2 decimal places and the SD to 3 decimal places; and so forth. Minimum and maximum will be displayed as reported in the source (raw) data. In addition, change from baseline values (visit value – baseline value) will be calculated at each time point and summarized descriptively. For categorical endpoints, descriptive summaries will include counts and percentages. Percentages will be presented to 1 decimal place unless otherwise specified; the denominator of percentages will be the number of subjects in the analysis set used unless specified otherwise.

All analyses and summary outputs will be generated using SAS® 9.4 or higher.

All summaries and figures will be provided by treatment group. Efficacy data are analyzed by the treatment the subject is randomized to. Safety data are analyzed by the actual treatment received, and this is determined by the highest strength of drug the subject received at any point of time during the study, irrespective of the number of days of exposure. For example, if a subject received placebo for 8 days and SAGE-217 20mg for 7 days, the actual treatment for the subject is SAGE-217 20mg. Unless specified otherwise, analyses using Safety Set will use actual treatment received; all other analyses will use randomized treatment.

All subject data, including those derived, to support tables and figures will be presented in the subject data listings. In general, the subject data listings will be sorted by subject number and assessment visit and date (and time, if applicable). The treatment will be identified for each subject.

For the purpose of all safety and efficacy analyses, baseline is defined as the last non-missing measurement prior to the first dose of study drug, unless stated otherwise. If the time of an assessment is collected, baseline will be the latest assessment prior to first dose of study drug administration time;

if the time of an assessment is not collected, the assessment on Day 1 is assumed to be prior to dosing if the protocol mentions that this assessment needs to be before dosing or it is collected as "pre-dose".

9.1.1 Study Day Definition

It is to be noted that the study drug is administered in the evening with food. The assessments at the clinic on Day 1 are hence before the first dose of study drug.

Study day will be defined as follows:

- The day of subject receiving the first dose of study drug is designated as Day 1.
- For visit days after Day 1, study day = visit date – Day 1 date + 1.
- For visit days prior to Day 1, study day = visit date – Day 1 date. Thus, study days for screening visit are negative numbers. There is no "Day 0".

9.1.2 Missing Data

All subjects will be used in the analyses, as per the analysis populations, using all non-missing data available. Efficacy analyses will use sensitivity analyses to assess the impact of missing data. Safety analyses will not impute missing data. Imputation of missing data in scoring of questionnaires is discussed in respective sections below. Handling of missing or incomplete dates have been discussed in Section 12.3, Appendix C.

9.2 Background Characteristics

All displays in this section will be presented under the treatment arms of placebo, SAGE-217 20mg, SAGE-217 30mg, SAGE-217 (20mg or 30mg) and Overall.

9.2.1 Subject Disposition

The analyses of subject disposition will use all subjects who provided written inform consent to the study.

The summaries of subject disposition will include the number of subjects who were screened, who were randomized, who received study drug, the number and percentage of subjects who completed the study, who are ongoing in the study (applicable only for the Day 42 Database Lock), who prematurely withdrew from the study, primary reasons for not completing the study, who completed treatment, who discontinued treatment prematurely, and primary reasons for discontinuing treatment. Study completion summary and treatment completion summary will be based on subjects who received study drug (Safety Set). Percentages will be calculated based on Safety Set. These data will be provided by randomized treatment groups. If a subject is rescreened because the subject has been a screen failure the first time, the status of the subject will be determined from the second screening. In the count of screened subjects, this subject will be counted only once.

A completer for the study is defined as one who completed the last follow up visit and is derived from the study conclusion CRF page with the completion question answered Yes. Note that in earlier versions of the protocol, this will be the completion of Day 42 while for later versions of the protocol this will be the completion of Day 182.

A subject is marked as completing the treatment if the prematurely discontinued question in the treatment discontinuation CRF page is answered No.

The number and percentage of subjects in each analysis set will be provided, using Safety Set as the denominator. Using Randomized Set, the reason for not being included in other analysis sets will be summarized.

The summary of subject disposition will also be provided using Modified Full Analysis Set.

A separate data listing will be provided for all subjects who prematurely discontinued treatment or prematurely withdrew from the study with reasons, number of days on study drug, etc., using Safety Set.

9.2.2 Protocol Deviations

Protocol deviations identified during site monitoring will be captured on eCRF and categorized by the study team as major and minor deviations, without any unblinding information. Protocol deviations, including but not limited to the use of prohibited medications, may be identified by blinded data review; these will be included in the list and categorized similarly. The major deviations are further categorized as major-efficacy, major-safety and major-GCP deviations. The major deviations will be summarized by type and by actual treatment received using Safety Set. The minor deviations will be included in the listing.

The deviations identified as major deviations will be reviewed by the study team prior to database lock in a blinded fashion to determine subjects in FAS to be excluded from the Per Protocol Set. In addition, Per Protocol Set will also exclude FAS subject satisfying any of the following conditions:

1. Study Drug Adherence (defined in Section 9.2.7) < 75% or less than 11 doses consumed
2. Inappropriate drug consumption (typically due to incorrect kit dispensation): if the subject consumed at any time during the study any study drug that the subject is not randomized to

The above criteria will be included in the major protocol deviation listing. In addition, a listing of inappropriate study drug consumption will be provided, which will include number of inappropriate doses consumed. Any violation of inclusion/exclusion criteria will be presented in a data listing using Randomized Set.

9.2.3 Demographics and Baseline Characteristics

The following analyses will use the Safety Set (using actual treatment received), the FAS and mFAS (using randomized treatment).

Demographic data (age, race, gender, ethnicity, employment status, highest education level, marital/civil status) and baseline characteristics, such as height, weight, and body mass index (BMI), will be summarized by treatment group (randomized for FAS and actual for Safety Set) and pooled treatment groups. Highest education level will be categorized in the summary tables as follows:

Less than or equal to12th grade, no diploma
12th grade diploma or GED
Some college but no degree
Associate degree
Bachelor's degree
Master's degree
Professional degree
Doctoral degree

Baseline subgroups will be summarized for the following categories (for definition of antidepressant, see Section 9.2.5):

- Race (Black or African American, White, Other)
- Gender (Female, Male)
- Age (18-24, 25-50, 51-65 years)

- Baseline antidepressant use (Yes, No)
- BMI (≤18.4, 18.5-24.9, 25-29.9, ≥30 kg/m2)
- Baseline HAM-D total score (<22, >=22 but <25, >=25)

A separate listing for subjects who were randomized under incorrect stratification of anti-depressant use at baseline will be provided. The correct use is determined from concomitant medication records with Anatomic Therapeutic Classification (ATC) level 3 = N06A, comparing the start/end dates of the medication versus the date of first dose of study drug. The stratum for randomization is determined from the eCRF entry of what was entered in IRT system during randomization.

Diagnostic labs are part of screening. A data listing using the Safety set will be provided. The following diagnostic screening test results will be included in this listing.

Diagnostic		
Serum	**Urine**	**Breathalyzer**
Hepatitis B	Drug screen: including amphetamines, barbiturates, benzodiazepines, cannabinoids, cocaine, opiates, phencyclidine	Alcohol
Hepatitis C	Female subjects that are not surgically sterile and do not meet the protocol-defined criteria for being post-menopausal: urine hCG	
Reflex HCV RNA		
HIV-1 and -2		
Female subjects that are not surgically sterile and do not meet the protocol-defined criteria for being post-menopausal: serum hCG		
Female Subjects, if menopause is suspected and not surgically sterile: FSH		

Abbreviations: FSH=follicle stimulating hormone; hCG=human chorionic gonadotropin; HCV = hepatitis C virus; HIV = human immunodeficiency virus

9.2.4 Medical/Surgical History

The following analyses will use the Safety Set.

The history related to MDD (date of initial diagnosis of MDD, antidepressant usage, information of depressive episodes, etc.) will be collected. Years since initial diagnosis of MDD, anti-depressant usage, and information of depressive episodes will be summarized. Years since initial diagnosis of MDD, days since start of current episode and years since start of first episode will be calculated using: First dose date of the study drug – date of interest. For imputation of incomplete dates in medical history, please see Section 12.3.3.

Medical/surgical history collected at screening will be coded using the Medical Dictionary for Regulatory Activities (MedDRA), Version 21.0 or higher. Medical/surgical history data will be summarized by system organ class (SOC) and preferred term (PT). A summary of medical/surgical history that are ongoing at the time of screening will be provided separately.

Subject history of psychiatric disorders and family psychiatric history will be summarized.

9.2.5 Prior and Concomitant Medications / Concomitant Procedures

The following analyses will use the Safety Set.

All medications taken and procedures undergone during the study will be recorded; in addition, psychotropic medications taken within 6 months prior to screening, and non-psychotropic medications taken within 30 days prior to screening will also be collected. Psychotropic and non-psychotropic medications are collected on separate CRF pages. All medications will be coded using World Health Organization-Drug dictionary (WHO-DD) March 2018 or later.

Medications will be presented according to whether they are being taken prior to and/or during the study (concomitant). Prior medications are defined as those taken prior to the initiation of the start of study drug. Concomitant medications are defined as those 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 concomitant. For imputation of missing concomitant medication dates, please refer to Appendix C, Section 12.3. Note that it is possible for a medication to be both 'prior' and 'concomitant'.

Concomitant medications will be further divided by usage period as follows (if time is missing, the date will be used for this algorithm):

- On-treatment concomitant medications are those that have been used any time from start of first dose to the last dose of study drug.
- Post-treatment concomitant medications are those that have been started after the last dose of study drug.

Prior and concomitant non-psychotropic medication use will be summarized by anatomical therapeutic chemical (ATC) level 1 and Standard Medication Name. Similar summary tables will be provided for psychotropic medications. The differentiation of non-psychotropic versus psychotropic medications are based on which CRF page the medication is collected on. Separate but similar summaries will be provided for concomitant medication use for on-treatment and post-treatment periods as defined above.

In addition, for prior and concomitant psychotropic medication separate summaries will be provided by ATC level 1 and ATC level 4.

Antidepressants that have been taken at the same dose for at least 60 days prior to the first dose of study drug are permitted if the subject intends to continue the stable dose through the follow-up period (Day 42). Initiation of new antidepressants or any other medications that may potentially have an impact on efficacy or safety endpoints is prohibited from at least the screening visit and through completion of the Day 42 assessments. Anti-depressant medications are identified by ATC3 level code of N06A. A summary of anti-depressant use at baseline and any change in these medications post-baseline (including the follow up period) will be provided.

Atypical antipsychotics will be identified by dictionary code of ATC level 3 equal to N05A, and will be listed separately. These medications are prohibited at entry in the Protocol Amendment 4 and beyond, but was permitted in earlier protocol versions as long as as these are taken in stable dose for at least 60 days prior to the first dose of study drug and through Day 42. These medications will include, but not limited to:

Aripiprazole
Asenapine
Brexpiprazole
Cariprazine
Clozapine
Iloperidone
Lurasidone
Olanzapine
Paliperidone
Pimavanserin
Quetiapine
Risperidone Ziprasidone

Concomitant procedures are recorded on a separate eCRF page; this will be presented in a listing by subject, and will not be summarized. The study day for the end date of the procedure will be provided, when a complete end date is available.

Prohibited medications are reviewed by the the medical monitor in the study team on an ongoing basis in blinded fashion; any medication identified as prohibited medication intake is captured in the protocol deviations list.

9.2.6 Study Drug Exposure

The following analyses will use the Safety Set.

Total drug exposure (in mg) is defined as the total study drug in mg for SAGE-217 that was taken during the study. Total drug exposure for subjects randomized to placebo is zero, unless the subject has taken SAGE-217 by mistake, in which case the total exposure comes from SAGE-217 exposure. If the patient skips the dose on any of the days, the dose taken is 0 mg.

Total exposure duration to study drug (in days) is defined as: Date of last dose – date of first dose + 1. Note that this does not exclude days when the dose has been missed.

Percent of the planned exposure received is defined as the total drug exposure, divided by planned exposure, times 100. For subjects who complete the treatment period, 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, and X=20 for subjects randomized to 20mg. For subjects who discontinue the treatment early, the planned exposure is (Last dose date – First dose date + 1), times X mg for subjects randomized to SAGE-217. For subjects randomized to placebo, this measure is not applicable.

Total drug exposure, total exposure duration and percent of the planned exposure received will be summarized descriptively. Number and percentage of subjects with less than 11 doses consumed will be provided.

9.2.7 Study Drug Adherence

The following analyses will use the Full Analysis Set as well as mFAS separately. Study drug adherence (%) is defined as the number of capsules taken, divided by the number of capsules planned to be taken, times 100.

The schedule of study drug is one capsule per day, so the number of days planned for study drug intake is the same as the number of capsules planned to be taken. The number of planned days for study drug intake is defined as follows:

1. If the subject discontinues treatment within Day 2 and Day 14 (both inclusive), the planned number of days is the last dose day of study drug.
2. If the subject does not discontinue treatment, the planned number of days is 14.

Study drug adherence will be summarized descriptively. Number and percentage of subjects with study drug adherence in categories - <75%, 75-100%, >100% - will be provided.

9.3 Efficacy Analysis

9.3.1 Definition of Efficacy Variables

The efficacy variables are defined as follows:

9.3.1.1 Hamilton Rating Scale for Depression (HAM-D)

The 17-item HAM-D will be used to rate the severity of depression in subjects already diagnosed as depressed. HAM-D is collected during the clinic visit on Days 1, 3, 8, 12, 15, 18, 21, 28, 35, 42, 70, 126, 182 .The 17-item HAM-D comprises of 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. Each item is scored in a range of 0 to 2 or 0 to 4, with higher scores indicating a greater degree of depression. The score for each item will be summed to compute a total score, which ranges from 0 to 52. If more than 3 individual items are missing a response, 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 mean of all other available item scores, or the maximum possible values for the missing responses, whichever is smaller, to calculate the HAM-D total score.

Four HAM-D subscale scores will be calculated will be calculated as the sum of the individual rating scores related to each subscale, divided by the total possible score within the subscale, multiplied by 100, and rounded to a whole number. If more than one item is missing or HAM-D total score is missing, the subscale score is left as missing; if one item on a particular subscale is missing, but has been imputed for the calculation of total score, the imputed value from total score calculation will be used in subscale score calculation for that item. Following table describes the subscale score calculation:

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.
Maier	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.

HAM-D response will be defined as having a 50% or greater reduction from baseline in HAM-D total score; only subjects who have a non-missing total score of HAM-D at baseline as well as the visit will be considered in HAM-D response evaluations. HAM-D remission will be defined as having a HAM-D total score of ≤7; if HAM-D total score is missing, remission will not be defined. For a sensitivity analysis the worst-case scenario imputation will be used, i.e. missing values for HAM-D response (remission) will be considered as "No response" ("No remission").

9.3.1.2 Clinical Global Impression – Improvement (CGI-I)

The Clinical Global Impression - Improvement (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. 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 on Days 3, 8, 12, 15, 18, 21, 28, 35, 42, 70, 126, 182. By definition, all CGI-I assessments are evaluated against baseline conditions. CGI-I response will be defined as having a CGI-I score of "very much improved" or "much improved." Missing CGI-I at the visit will not be evaluated for response.

9.3.1.3 Clinical Global Impression – Severity (CGI-S)

The Clinical Global Impression - Severity (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. CGI-S is collected at Screening day, Days 1, 3, 8, 12, 15, 18, 21, 28, 35, 42, 70, 126, and 182.

9.3.1.4 Hamilton Anxiety Rating Scale (HAM-A)

The 14-item HAM-A will be used to rate the severity of symptoms of anxiety. HAM-A is collected during the clinic visit at Days 1, 8, 15, 18, 28, 42. 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 range of 0 to 56, where <17 indicates mild severity, 18 to 24, mild to moderate severity, and 25 to 30, moderate to severe severity. The HAM-A total score will be calculated as the sum of the 14 individual item scores. If more than 3 individual items are missing a response, the HAM-A 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 mean of all other available item scores, or the maximum possible values for the missing responses, whichever is smaller, to calculate the HAM-A total score.

9.3.1.5 Montgomery-Åsberg Depression Rating Scale (MADRS)

The MADRS is a 10-item diagnostic questionnaire used to measure the severity of depressive episodes in subjects with mood disorders. MADRS is collected at Screening, during the clinic visit on Days 1, 3, 8, 12, 15, 18, 21, 28, 35, and 42.

Each MADRS item ranges from 0 to 6; higher MADRS scores indicate more severe depression. The MADRS total score will be calculated as the sum of the 10 individual item scores, which ranges from 0 to 60. If more than two individual items are missing, the MADRS total score will not be calculated and will be left as missing. If less than or equal to two individual item scores are missing, the missing item scores will be imputed by the mean of all other available item scores, or the maximum possible values for the missing responses, whichever is smaller, to calculate the MADRS total score.

MADRS response will be defined as having a 50% or greater reduction from baseline in MADRS total score; only subjects who have a non-missing total score of MADRS at baseline as well as the visit will be considered in MADRS response evaluations. MADRS remission will be defined as having a MADRS total score of ≤10; if MADRS total score is missing, remission will not be defined.

9.3.1.6 Insomnia Severity Index (ISI)

The ISI is a validated, 7-item questionnaire designed to assess the nature, severity, and impact of insomnia. It is collected during the clinic visit on Days 1, 8, 15, 18, 21, 28, 42, 70, 126, and 182. The ISI uses a 5-point Likert Scale to measure various aspects of insomnia severity (0 = none, 1 = mild, 2 = moderate; 3 = severe; 4 = very severe), satisfaction with current sleep pattern (0 = very satisfied, 1 = satisfied, 2 = moderately satisfied, 3 = dissatisfied, 4 = very dissatisfied), and various aspects of the impact of insomnia on daily functioning (0 = not at all, 1 = a little, 2 = somewhat, 3 = much, 4 = very much). The total score is derived as the sum of item scores. If more than 1 individual item is missing, the ISI total score will not be calculated and will be left as missing. If 1 individual item score is missing, the missing item score will be imputed by the mean of all other available item scores to calculate the ISI total score. Total score is categorized as: 0 to 7 = “no clinically significant insomnia”, 8 to 14 = subthreshold insomnia”, 15 to 21 = “clinical insomnia (moderate severity) and 22 to 28 = “clinical insomnia (severe).” Missing ISI total score will not be categorized.

9.3.1.7 Core Consensus Sleep Diary

The take-home subject sleep diary assessment will be administered using an eDiary solution. The eDiary will be captured using either a provisioned smartphone device or bring-your-own-device (BYOD) solution, depending on the subject’s preference. The sleep diary will be collected at least 7 days prior to Day 1 and then daily through Day 28. The sleep diary is to be completed each day in the morning, preferably in the morning after awakening from sleep, regarding the last night’s sleep.

Subjective sleep parameters will be derived for study days 8, 15, 21, 28 using the mean of non-missing assessments between the previous visit day (exclusive) and the current visit day (inclusive). The baseline subjective sleep parameters will be calculated as the mean of 7 days immediately preceding day 1. In addition, Day 18 visit values will be calculated using the days after Day 15 visit until Day 18 visit, but this timepoint will only be displayed in summary and listing, not used for analysis model.

Parameters to be analyzed are as follows:

sTST (in minutes) = Total sleep time = Time of final awakening - (Time when tried to sleep + time taken to fall asleep) – time of being awake after sleep onset
sSL (in minutes) = Sleep Latency
sWASO (in minutes) = Wake after Sleep Onset = sWASO
sNAW = Number of awakenings = sNAW
sSQ = Sleep Quality

The date/time stamp on the day's entry is compared to clinic visit date to identify records that go into the calculation of a visit assessment. In case of duplicate assessment on the same day, the first assessment will be counted toward the mean, but all assessments will be listed. The times entered in the diary by the subject, e.g. time when went to bed, time when tried to fall asleep, etc. need to be associated with a date in order to make meaningful comparisons or to be used in calculations. The following algorithm will be followed for this purpose: If Time to Bed is AM, then all time questions are associated with stamped date. If Time to Bed is PM and Time out of Bed is PM, then all time questions are associated with date one day before the stamped date. Otherwise, when Time to Bed is PM, all times in diary marked as PM are associated with day before the stamped date, and all times in diary marked as AM are associated with stamped date.

Since the data is subject-entered and not monitored, a record may be invalid when it fails to make logical sense, particularly in calculation of sTTB or sTST, A record is considered invalid if either of the following holds true (use date/time combination to do these comparisons):

1. Any of the eight response or date stamp is missing.
2. sTST > sTTB (Total time in bed, in minutes, calculated as Time out of bed – Time to bed)
3. sTST < 0
4. sTTB <= 0
5. Time to bed > Time when tried to sleep
6. Time to bed > Time of final awakening
7. Time when tried to sleep > Time of final awakening
8. Time tried to sleep > Time out of bed
9. Time of final awakening > Time out of bed

An invalid record will be dropped from analysis of sTST, but not from the analysis of other sleep diary endpoints. The sleep diary data will not be windowed.

Sleep quality (sSQ) response will be defined as having a sSQ score of "very good" or "good." Missing sSQ at the visit will not be evaluated for response.

9.3.1.8 Short Form-36 Version 2 (SF-36v2)

The Medical Outcomes Study Short Form-36 version 2 (SF-36v2) is a 36-item measure of health status that has undergone validation in many different disease states. This is collected during the clinical visits at Screening, on Days 1, 8, 15, 28, 42, 70, 126, and 182.

The SF-36 covers 8 health dimensions including 4 physical health status domains (physical functioning, role participation with physical health problems [role-physical], bodily pain, and general health) and 4 mental health status domains (vitality, social functioning, role participation with emotional health problems [role-emotional], and mental health). In addition, two component summary scores, physical component summary and mental component summary (MCS), are produced by taking a weighted linear combination of the 8 individual domains. There is also an Utility Index score (Release 2) that is available in SF-36 scale. Higher scores indicate a better state of health.

The scoring of this questionnaire is proprietary to Optum Incorporated; it involves using current norms of relevant populations. The raw data will be provided to Optum, and they will return the validated, quality-checked derived scores for each subject at each assessment, which will be used for analyses.

9.3.1.9 Patient Health Questionnaire (PHQ-9)

The PHQ-9 is a 9-item subject-rated depressive symptom severity scale. It is collected during the clinic visit on Days 1, 8, 15, 21, 42, 70, 126, and 182. Scoring is based on responses to specific questions, as follows: 0=not at all; 1=several days; 2=more than half the days; and 3=nearly every day.

The PHQ-9 total score will be calculated as the sum of the 9 individual item scores. If more than 1 individual item is missing, the PHQ-9 total score will not be calculated and will be left as missing. If 1 individual item score is missing, the missing item score will be imputed by the mean of all other available item scores to calculate the PHQ-9 total score. The PHQ-9 total score will be categorized as follows: 1 to 4=minimal depression, 5 to 9=mild depression, 10 to 14=moderate depression, 15 to 19=moderately severe depression; and 20 to 27=severe depression.

9.3.2 Visit Windows

The scheduled visits will not be windowed and will be used at nominal visit value for analysis purposes. The unscheduled, end-of-treatment (EOT) and early termination (ET) visit will be mapped to a scheduled visit for analysis. For unscheduled visits that happens on or before EOT visit date (including EOT visit) will be mapped using the date of collection/assessment and Day 1 first dose date as a basis to determine study day and then study day will be mapped to the intended visit according to the visit windows specified in the table below. Unscheduled visits after EOT visit date, including ET visit, will be windowed using relative days since last dose date; the mapping will follow the table below. In order to accommodate as much data as possible into analysis, these windows have been widened compared to protocol-specified operational window, to have no gap between them; these windows are used for analysis purposes only. Note that Day 18 data will not be windowed; it will be presented in summaries as nominal visit values, and will not be included in the modeling or in figures.

Once analysis visit windows are assigned, all visits, including scheduled visits, unscheduled visits, and EOT/ET visits will be eligible for being flagged as the “analyzed record” within the analysis window; 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 the data from the scheduled visit is available, then the scheduled visit data will be used.
- If there is no data from the scheduled visit is available, the data closest to the scheduled study day for that window will be used.
    - If there is a tie between the data in the number of days before and after the scheduled day, the later data will be used.

The summary by visit will use the "analyzed records" only – at most one per subject. The data not flagged as the "analyzed record" will be included in listings. An unscheduled visit that does not fall under any analysis window (e.g. in case one is available after Day 182) will remain in the database,and will be included in the listings.

Table 1 displays windows for efficacy analysis.

Table 1: Visit Windows for Efficacy Analysis

Scheduled Visit (+/1 window days) in protocol	Target Study Day	Study Day Window for Visit in Analysis
Screening	Day -1	Days (-28) to (-1)
First Dose	Day 1 (pre-dose)	Day 1 (pre-dose)
Day 3 (±1 day)	Day 3	Day 2 - Day 5
Day 8 (±1 day)	Day 8	Day 6 - Day 9
Day 12 (±1 day)	Day 12	Day 10 - Day 13
Day 15 (±1 day)	Day 15	Day 14 - Day 17
Day 21 / last dose + 7 days (±1 day)	Day 21 (last dose date + 7 days)	Day 18 - Day 23 (last dose date +4 days, +9 days)
Day 28 / last dose + 14 days (±3 day)	Day 28 (last dose date + 14 days)	Day 24 - Day 31 (last dose date +10 days, +17 days)
Day 35 / last dose + 21 days (±3 days)	Day 35 (last dose date + 21 days)	Day 32 – Day 38 (last dose date +18 days, +24 days)
Day 42 / last dose + 28 days (±3 days)	Day 42 (last dose date + 28 days)	Day 39 – Day 45 (last dose date +25 days, +31 days)
Day 70 / last dose + 56 days (± 7 days)	Day 70 (last dose date + 56 days)	Day 60 – Day 80 (last dose date + 46 days, +66 days)
Day 126 / last dose + 112 days (± 7 days)	Day 126 (last dose date + 112 days)	Day 116 – Day 136 (last dose date + 102 days, + 122 days)
Day 182 / last dose + 168 days (± 7 days)	Day 182 (last dose date + 168 days)	Day 172 – Day 192 (last dose date + 158 days, +178 days)

Note: Parenthesized study day and study day window are for unscheduled visits, EOT and ET for subjects who have discontinued treatment prematurely and such visit date is >=4 days from the last dose of study drug intake (i.e. visit date – last dose date + 1 >4).

9.3.3 Analysis of Efficacy Variable(s)

The modified FAS will be used for all efficacy summary tables, since this corresponds to the targeted patient population (see Section 7.1). Corresponding analysis in FAS will be provided for supportive purposes only. Subjects will be analyzed according to randomized treatment.

The following efficacy endpoints will be summarized descriptively by scheduled assessment time point:

- HAM-D total score – observed, change from baseline, percent change from baseline
- HAM-D subscale scores – observed, change from baseline, percent change from baseline
- HAM-D individual item score -– observed, change from baseline
- HAM-D response – missing response not accounted
- HAM-D response – missing response counted as No response
- HAM-D remission – missing remission not accounted
- HAM-D remission – missing remission counted as No remission
- CGI-I score -observed
- CGI-I response
- CGI-S score – observed and change from baseline
- HAM-A total score – observed, change from baseline, percent change from baseline
- HAM-A individual item score -– observed, change from baseline
- MADRS total score – observed, change from baseline, percent change from baseline
- MADRS individual item score – observed, change from baseline
- MADRS response – missing response not accounted
- MADRS remission – missing remission not accounted
- ISI total score– observed (including categories), change from baseline (including any shift from baseline in categories)
- Sleep endpoints (eg, subjective sleep latency (sSL), subjective total sleep time (sTST), subjective wake after sleep onset (sWASO) – observed, change from baseline, percent change from baseline
- SF-36v2 domain/component score– observed, change from baseline, percent change from baseline

  In addition to the three randomized treatment arms in the study, the SF-36 data will also be summarized for SAGE-217 irrespective of dose (20mg and 30mg dose arms combined in one).
- PHQ-9 score – observed (including categories), change from baseline (including any shift from baseline in categories)
- PWC-20 – observed, change from baseline

The HAM-D change from baseline in total score, response and remission will also be presented by the following subgroups:

- Antidepressant use at baseline: yes, no
- Age group: 18-24, 25-50, 51-65 years

- Gender: Male, Female
- Race: White, Black or African American, Other
- BMI (≤18.4, 18.5-24.9, 25-29.9, ≥30 kg/m2)
- Baseline HAM-D total score: <22, >=22 but <25, >=25

In addition, post-baseline percentage improvement in HAM-D total score will be presented in histogram over scheduled visits by treatment group under the following categories: < 0% (worsened), >= 0% but < 25%, >=25% but < 50%, >= 50% but <75%, >=75%. Post baseline HAM-D remission will also be presented in histogram with HAM-D total score over time: <=7, >7 but <=15, >15. Supporting data will be presented in summary tables.

Bar charts over scheduled visits by treatment for HAM-D response, HAM-D remission and CGI-I response will be provided.

As part of supportive analyses, the HAM-D total scores, MADRS total scores, HAM-D response and HAM-D remission will be summarized for Per Protocol Set as well as FAS separately. Line plots, histograms and bar charts will be provided for FAS when applicable.

9.3.3.1 Mixed Effects Model for Repeated Measures

The estimand for the primary efficacy analysis is defined as follows:

1) The target population is adult subjects with a diagnosis of major depressive disorder and within a current depressive episode of moderate severity (HAM-D total score >=22).
2) The variable of interest is the change from baseline in HAM-D total score at Day 15.
3) The population summary level deals with the difference between SAGE-217 and placebo treatments in mean change from baseline in HAM-D total score at Day 15.
4) The intercurrent events could be:
    a. the premature discontinuation of treatment for any reason, thus not having a Day 15 HAM-D total score available. This will be dealt by a sensitivity analysis using multiple imputation technique as described below.
    b. Certain medications including, but not limited to, new antidepressants or benzodiazepines are prohibited in the protocol until Day 42 follow-up; however, the treatment policy strategy dictates that the results following these prohibited medication use will not be manipulated, but will rather be used ‘as is’ in analysis. Please note that the protocol does not specify any rescue process, hence there is no rescue medication.

Data from SAGE-217 30mg group versus placebo group will be analyzed using a mixed effects model for repeated measures (MMRM); the model will include treatment (SAGE-217 30 mg, SAGE-217 20 mg or placebo), baseline HAM-D total score, anti-depressant use at baseline (Yes or No), assessment time point (excluding Day 18), and time point-by-treatment as explanatory variables. All explanatory variables will be treated as fixed effects. All post-baseline time points will be included in the model, except Day 18. The main comparison will be between SAGE-217 and placebo at the 15-day time point. Model-based point estimates i.e., least squares [LS] mean, is used as the test statistic and will be reported where applicable along with 95% confidence intervals, and p-values. An unstructured (UN) covariance structure will be used to model the within-subject errors. 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. The sandwich estimator for the variance covariance matrix will be derived, using the

EMPIRICAL option in the PROC MIXED statement in SAS. The p-value will be interpreted at 5% level of significance.

If the comparison of SAGE-217 30mg versus placebo is significant at 0.05 level, other hypotheses testing will proceed with multiplicity adjustment, as described in Section 9.3.3.5.

Similar to those methods described above for the primary endpoint, a MMRM will be used for the analysis of the change from baseline in other time points in HAM-D total score, all time points in HAM-D subscale scores (Core Subscale score, Anxiety Subscale score, Bech-6 Subscale score and Maier Subscale score), HAM-D individual item scores, CGI-S score, HAM-A total score, MADRS total score, MADRS individual item scores, ISI score, sleep endpoints (eg, sleep onset latency (sSL), total sleep time (sTST), wake after sleep onset (sWASO). Number of awakenings (sNAW)), SF-36v2 Domain/Component Score, and PHQ-9 total score. The explanatory variable of baseline value will use the respective parameters being analyzed. The analyses will also be provided for change from baseline in HAM-D total score within each baseline subgroup level separately.

For each model, the comparison of interest will be between SAGE-217 and matching placebo at the 15-day time point, with 30mg comparison followed by 20mg comparison, according to the multiplicity adjustment, as described in Section 9.3.3.5. However, model-based point estimates (i.e., LS means), 95% confidence intervals, and p-values will be reported for all time points.

Line plot of model-based LS Mean and standard error (SE) over time will be prepared for change from baseline in HAM-D total score, HAM-D subscale scores, CGI-S score, HAM-A total score, MADRS total score, ISI score, sleep endpoints (total sleep time (sTST), sleep latency (sSL), wake after sleep onset (sWASO), number of awakenings (sNAW), and PHQ-9 score. Forest plot for subgroup analysis for change from baseline in HAM-D total score at Day 15 – LS means, confidence interval – will be provided. In addition, a Forest plot for individual items in HAM-D and for subscales of HAM-D will also be provided. For SF-36, LS means of change from baseline at Day 15 will be provided in bar chart for each domain/component.

Summary of HAM-D total score along with model-based estimates and line plot will be provided for Per Protocol Set as well as for FAS as a supportive analysis for the primary endpoint.

9.3.3.2 Generalized estimating equation (GEE) models

Generalized estimating equation (GEE) methods will be used for the analysis of HAM-D response and HAM-D remission. GEE models will include terms for treatment (SAGE-217 30mg, SAGE-217 20mg, Placebo), baseline HAM-D score, antidepressant use at baseline (Yes or No), assessment time point, and time point-by-treatment as explanatory variables. The comparison of interest will be the difference between SAGE-217 and matching placebo at the 15-day time point – 30mg comparison, followed by 20mg comparison, according to the multiplicity adjustment, as described in Section 9.3.3.5. Model-based point estimates (i.e., odds ratios), 95% confidence intervals, and p-values will be reported.

A GEE method will also be used for the analysis of MADRS response and remission including terms for treatment, baseline MADRS score, antidepressant use at baseline, assessment time point, and time point-by-treatment as explanatory variables.

A GEE method will also be used for the analysis of CGI-I response including terms for treatment, baseline CGI-S score, antidepressant use at baseline, assessment time point, and time point-by-treatment as explanatory variables.

A GEE model will be used for analysis of sleep quality (sSQ) response including terms for treatment, baseline sSQ, antidepressant use at baseline, assessment time point, and time point-by-treatment as explanatory variables.

9.3.3.3 Sensitivity Analysis

A sensitivity analysis will be used to investigate the impact of missing data if ≥ 5% of subjects have missing data in primary efficacy endpoint assessment (i.e., HAM-D total score at Day 15).

Imputation based on study withdrawal reason will be used. The missing change from baseline in HAM-D total scores will be imputed using multiple imputation methods. The MMRM model will use the imputed dataset (all observed and imputed values included) to estimate the treatment difference. The modified FAS will be used for sensitivity analyses, with FAS as supportive analyses. Sample SAS codes for MI imputation is provided in Section 12.2, Appendix B.

Imputation distribution:
The imputation distribution for the missing change from baseline in HAM-D total score at each timepoint will be assumed to have a normal distribution. All randomized 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 X
- Missing category 1: Subject discontinued due to adverse events, physician decision, protocol deviation, non-compliance with study drug, , sponsor decision or other, and is missing change from baseline in HAM-D total score at Day X
- Missing category 2: Subject discontinued due to pregnancy, lost to follow-up, subject decision, withdrawal by subject, and is missing change from baseline in HAM-D total score at Day X, or subject completed study but is missing change from baseline in HAM-D total score at Day X

Imputation algorithm:
Missing values of change from baseline in HAM-D total score 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 $N(\mu, \sigma^2)$, where μ is the mean of the non-missing change from baseline in HAM-D total score for placebo group for Day X, and σ^2 is the sample variance estimated using the non-missing change. This represents a conservative approach since higher values of change from baseline represents worse outcome, and placebo is supposed to provide a higher value.
- **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 for subjects within the same treatment group at Day X, and σ^2 is the sample variance estimated using the non-missing change.

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 FAS subjects will have non-missing change from baseline in HAM-D total score at all scheduled assessment time points.
- Repeat the process K (K=20) times, using the procedure described above to form K imputed complete datasets.
- Fit the MMRM model including treatment, baseline antidepressant use, and baseline HAM-D total score, to each imputed dataset, to estimate the treatment effect and its variance.
- Combine the results from the K imputed datasets using the SAS procedure MIANALYZE, to derive the MI estimator.

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

It has been shown that the statistic

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

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

9.3.3.4 Analysis of Time to First HAM-D Response/Remission

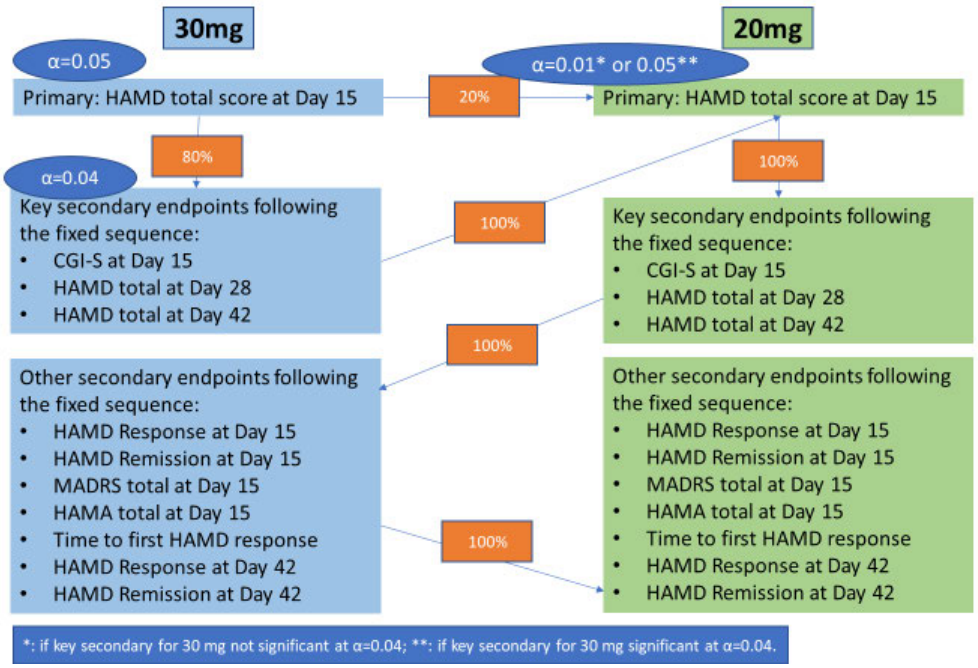
Using the modified FAS, Kaplan-Meier (KM) survival curve will be provided for time to first HAM-D response; the median time to first response will be estimated from KM analysis. A subject will be censored at the subject's last day of HAM-D evaluation in the database if the subject did not have a response. Similar analysis will be done for first HAM-D remission.

These analyses will be provided on FAS for supportive purposes only.

9.3.3.5 Multiplicity Adjustment versus Key Secondary Endpoints

Multiplicity adjustment to statistical testing of hypotheses of secondary endpoints are done through dividing the alpha of 0.05 for primary testing into 30mg and 20mg in 80% and 20%, then proceeding sequentially with pre-identified key secondary endpoints. Figure 1 below explains the strategy. A secondary endpoint not included below is not adjusted for multiplicity, and hence will be interpreted with nominal p-value.

Figure 1: Multiplicity Adjustment Strategy for Testing Key Secondary Endpoints



9.4 Safety Analysis

The primary objective is to evaluate the safety and tolerability of SAGE-217 as assessed by the incidence and severity of adverse events; changes from baseline in clinical laboratory measures, vital signs, and ECGs; and suicidal ideation and behavior using the C-SSRS. Potential withdrawal symptoms after discontinuation of SAGE-217 will be assessed using the PWC-20. Safety analyses will be conducted using the Safety Set, unless specified otherwise. The data will be presented by the actual treatment received rather than the treatment to which the subject has been randomized; for definition of actual treatment assignment, please see Section 9.1. In addition to the three treatment arms in the study, the safety data will also be presented for SAGE-217 irrespective of dose (20mg and 30mg dose arms combined in one).

The safety endpoints evaluated at scheduled visits are taken as done in nominal visit, without any windowing. If a value is available for a nominal scheduled visit, that value will be used in summary by visit. Unscheduled visits, EOT and ET visits will be windowed using the same window days outlined in Table 1 for efficacy endpoints. If scheduled visit value is not available, a value from the specific visit window will be included in summary, the choice of the record following the same rule as described in Section 9.3.2.

Last value on treatment and Last value on study will be included in the summaries whenever indicated in the relevant sections below. Last value on treatment is defined as the last post-baseline value between first dose of study drug (exclusive) and up to last dose of study drug + 7 days (inclusive). Last value on study is defined as the last post-baseline value after the first dose of study drug.

The safety endpoints and variables considered in the summary tables for this study are summarized in Table 2.

Table 2: Safety endpoints and variables in the summary tables

Safety Evaluation	Incidence	Observed Value	Change from Baseline	Abnormality/Clinical Significance (CS)	Potentially Clinical Significance (PCS)

Safety Evaluation	Incidence	Observed Value	Change from Baseline	Abnormality/Clinical Significance (CS)	Potentially Clinical Significance (PCS)
AEs	X				
Labs		X	X	Z	X
ECGs		X	X	Z	X
Vital Signs		X	X		X
C-SSRS	X	X	X		
PWC-20		X	X		

Note: PCS criteria are outlined in sections 9.4.2-9.4.4

X = to be summarized in tables
Z = to be presented in listings only

9.4.1 Adverse Events

Adverse events (AEs) are collected starting at the time of informed consent and throughout the duration of the subject's participation in the study. A treatment-emergent adverse event (TEAE) is defined as an adverse event with onset after the start of study drug. The TEAEs will be further categorized by the phase of occurrence as follows:

Adverse events are assigned an AE period based on the onset date/time. AE periods are defined as follows:

- Pre-treatment AE: AE onset date before first study drug dosing date
- TEAE: AE onset date/time on or after first study drug dose date/time (If an AE start date same as study drug first dose date, but no time either in AE start or treatment start, then consider this AE to be in treatment period TEAE.)
- Treatment Period AE: AE onset date/time on or after first study drug dose date/time and on or before study drug last dose date + 1 day (Note that time does not matter for the end of this period.)
- Follow-up period TEAE: AE onset date after study drug last dose date +1 day and on or before study drug last dose date + 28 days (Typically, Day 16 through Day 42 – time does not matter)
- Extended Follow-up Period TEAE: AE onset date after study drug last dose date + 28 days
- Double-blind Period TEAE: All AEs that are flagged as either Treatment Period or Follow-up Period TEAE

If the date of an adverse event is incomplete and an unambiguous determination could not be made with respect to its onset time versus the first dose of study drug and/or last dose of study drug, the adverse event will be assumed to be a TEAE and a treatment period AE. For imputation of missing AE dates, please refer to Appendix C, Section 12.3.1.

All adverse events will be coded using MedDRA version 21.0 or higher.

An overview summary table of TEAEs will present the number and percentage of subjects as well as the number of events for the following:

- TEAE Double-Blind Period TEAE
  - Treatment Period TEAE
  - Follow-up Period TEAE

    - Extended Follow-Up Period TEAE
- TEAEs by maximum severity (severe>moderate>mild)
    - Double-Blind Period TEAE
        - Treatment Period TEAE
        - Follow-up Period TEAE
    - Extended Follow-Up Period TEAE
- TEAE leading to discontinuation of study drug
- TEAE leading to withdrawal from the study
    - Double-Blind Period TEAE
        - Treatment Period TEAE
        - Follow-up Period TEAE
    - Extended Follow-Up Period TEAE
- Serious Adverse Event (SAE)
    - Double-Blind Period Serious AE
        - Treatment Period Serious AE
        - Follow-up Period Serious AE
    - Extended Follow-Up Period TEAE
- Death
    - Double-Blind Period
        - Treatment Period
        - Follow-up Period
    - Extended Follow-Up

Incidence of TEAEs in following categories will be provided by SOC and PT. A subject is counted only once under each SOC and PT in case of multiple occurrences of the same AE. These tables will be sorted by decreasing frequency in SAGE-217 30mg+20mg group, then in 30mg group, then in 20mg group, then in placebo group, then alphabetically.

- TEAE
- Double-Blind period TEAE
- Treatment Period TEAE
- Follow-up Period TEAE
- Extended Follow-up Period TEAE
- Double-Blind Period TEAEs by maximum Severity
- Double-Blind Period TEAEs by relationship

- Serious TEAEs
- Double-Blind Period Serious TEAEs
- Treatment Period Serious TEAE
- Follow-up Period Serious TEAE
- Extended Follow-up Period Serious TEAE
- Treatment Period TEAEs leading to discontinuation of study drug
- TEAEs leading to withdrawal from the study
- Double-Blind Period TEAEs leading to withdrawal from the study

Listing of AEs with onset prior to first dose of study drug will be provided. All listings on TEAEs will provide the period designation for each AE.

A summary of most common study period TEAE just by preferred term where the incidence is more than 2% in any treatment group will be provided, sorted by decreasing frequency first by SAGE-217 30mg+20mg group, then by 30mg, then by 20mg, then by placebo, then alphabetically.

For maximum severity, subjects will be counted only once within each SOC and PT at the maximum severity in the following order: severe> moderate> mild; an AE with missing severity will be omitted from severity presentation. For relationship to study drug, 'related' is defined as relationship being "possible" or "probable" or missing. A subject will be counted only once within each SOC and PT at the strongest relationship to study drug in the following order: related > not related. The incidences will be presented by descending frequency of SOC and then, within a SOC, by descending frequency of PT based on the subject count, and in alphabetical order of PT if the incidence within a PT is a tie. Adverse events with onset before the first dose of study drug will be provided in a separate listing. Separate data listing for deaths and non-fatal SAEs will be provided.

In addition, double-blind period TEAE summary by SOC/PT will also be presented by the following subgroups:

- Anti-depressant use at baseline: yes, no
- Age group: 18-24, 25-50, 51-65 years
- Gender: Male, Female
- Race: White, Black or African American, Other
- BMI (≤18.4, 18.5-24.9, 25-29.9, ≥30 kg/m2)
- Baseline HAM-D total score: <22, >=22 but <25, >=25

9.4.2 Clinical Laboratory

The clinical laboratory tests to be performed for monitoring of safety are listed in Table 3. They are collected on screening day, days 1, 8, 15, 21, 28, 42, 70, 126, 182.

Table 1: Clinical Laboratory Tests

Hematology	Serum Chemistry	Urinalysis	Coagulation
Red blood cell count Hemoglobin Hematocrit	Alanine aminotransferase Albumin Alkaline phosphatase	pH Specific gravity Protein	Activated partial Thromboplastin time

White blood cell count with differential Platelet count Red Blood Cell Indices (MCV, MCH, MCHC) Reflex to Red blood cell morphology if indices are abnormal	Aspartate aminotransferase Total bilirubin Direct bilirubin Indirect bilirubin Total protein Creatinine Blood urea nitrogen Creatine kinase Gamma-glutamyl transferase Potassium Sodium Lactate dehydrogenase Glucose Chloride Bicarbonate Calcium Phosphorus Triglycerides Thyroid stimulating hormone (TSH) Reflex to free T3/T4 if TSH is abnormal	Glucose Red blood cell Nitrite Leukocyte esterase Ketones Bilirubin Urobilinogen	Prothrombin time International normalized ratio

All parameters will be converted to consistent units according to the International System of Units (SI) before presentation.

For the laboratory results that is "< or = *x*", where *x* is a number as collected in the data, the numeric part of the result will be used in the calculation in the summary tables. The same is true if the result is presented as BLQ and a LLOQ value is provided – LLOQ value will be used for calculation in the summary tables. The actual results as collected will be displayed in the listings.

Summary tables on lab parameters will include descriptive statistics for the observed values and changes from baseline by scheduled assessment timepoint in routine parameters in hematology, serum chemistry, coagulation and quantitative urinalysis test results. It will also include the summary of last post-baseline values on treatment and on study. The parameter values, which are produced only if another parameter is abnormal, will be included in data listings, but not summarized.

If a normal range is provided for the parameter, out-of-range values will be flagged as low or high, where applicable, in the subject data listings. A shift table for these parameters from baseline to each scheduled assessment time point will be provided. This table will also include the shift from normal at baseline to high or low at any time during treatment (on or after first dose, on or before last dose + 1 day), any time post-baseline during the study (on or after first dose) and at the end of study (i.e. the last available value in the database). Qualitative urinalysis parameters will be summarized descriptively.

The number and percentage of subjects with PCS values will be provided in separate displays in hematology, serum chemistry, liver function tests and urinalysis tests provided for such occurrence any time post-baseline (irrespective of whether it happens in scheduled or unscheduled assessments). Potentially clinically significant values will be identified for specific laboratory parameters as outlined in the following table.

Laboratory Parameter	Gender	Units	Criteria for PCS Values (Observed values)	
			High	Low
Hematology				
Hemoglobin	Male	g/L	>185	<115
	Female	g/L	>170	<100
Hematocrit	Male	Fraction of 1	>0.55	<0.385
	Female	Fraction of 1	>0.49	<0.345
Platelet count		10^9/L	>600	<125
White blood cell		10^9/L	>15	<2.5
Basophils		10^9/L	>0.5	NA
Eosinophils		10^9/L	>1.5	NA
Neutrophils		10^9/L	NA	<1.5
Lymphocytes		10^9/L	>6.0	<0.5
Monocytes		10^9/L	>1.4	NA
Serum Chemistry				
Albumin		g/L	>70	<28
Blood urea nitrogen		mmol/L	>10.71	NA
Calcium		mmol/L	>2.75	<2.0
Chloride		mmol/L	>120	<90
Creatinine		mmol/L	>3xULN or >3x Baseline	
Gamma Glutamyl Transferase			>3xULN	
Glucose		mmol/L	>13.9	<2.8
Sodium		mmol/L	>150	<132
Potassium		mmol/L	>5.4	<3.3
Protein		g/L		<45
Bicarbonate		mmol/L	>34	<18
Chloride		mmol/L	>120	<90
Phosphorus		mmol/L	>1.94	<0.61
Liver Function Tests (LFT)				
Bilirubin		µmol/L	>2xULN	NA
Aspartate Aminotransferase		U/L	>3xULN	NA
Alanine Aminotransferase		U/L	>3xULN	NA
Alkaline Phosphatase		U/L	>1.5xULN	NA

Liver function tests will be monitored closely for potentially clinically significant values, and will be summarized for occurrence any time post-baseline for the following parameters for these PCS threshold (for condition involving more than one parameters, the results need to be from the same timepoint):

Alanine Aminotransferase: >3xULN, >5xULN, >10xULN
Aspartate Aminotransferase: >3xULN, >5xULN, >10xULN
Alanine Aminotransferase or Aspartate Aminotransferase: >3xULN, >5xULN, >10xULN
Alkaline Phosphatase: >1.5xULN, >2xULN
Total Bilirubin: >1.5xULN, >2xULN
Total Bilirubin > 2xULN **AND** (Alanine Aminotransferase or Aspartate Aminotransferase >3xULN)

Total Bilirubin >2xULN **AND** Alkaline Phosphatase >2xULN **AND** (Alanine Aminotransferase or Aspartate Aminotransferase >3xULN)

Any lab results considered clinically significant by the investigator will be captured as adverse events, hence will show up in AE displays.

Pregnancy test results will be listed but not summarized.

9.4.3 Vital Signs

Vitals for the following parameters - respiratory rate (breaths/minute), oral temperature (degrees C), supine heart rate (beats/minute), supine systolic blood pressure (mmHg), supine diastolic blood pressure (mmHg), standing heart rate (beats/minute), standing systolic blood pressure (mmHg), standing diastolic blood pressure (mmHg), – are collected at screening, days 1, 3, 8, 12, 15, 18, 21, 28, 35, 42, 70, 126, and 182. Descriptive summaries of observed values and changes from baseline will be provided for vital sign parameters - by scheduled assessment time point. It will also include the summary of last values on treatment and on study assessments. Additionally, the number and percentage of subjects with PCS and PCSC values will be summarized for such occurrence any time post-baseline. Potentially clinically significant values will be identified for vital sign parameters as outlined in the following table.

Vital Sign	**Units**	**Criteria for PCS Values (Observed values)**		**Criteria for PCSC values (Change from Baseline values)**	
		High	Low	Increase	Decrease
Heart rate (supine and standing)	Beats/min	>120	<40	NA	NA
Systolic Blood Pressure (supine and standing)	mmHg	>180	<90	≥30	≥30
Diastolic Blood pressure (supine and standing)	mmHg	>110	<50	≥20	≥20
Supine - Standing Systolic Blood Pressure	mmHg	≥20			
Supine – Standing Diastolic Blood Pressure	mmHg	≥10			
Orthostatic hypotension: supine – standing SBP and DBP	mmHg	SBP >= 20 and DBP >= 10			
	mmHg	SBP >= 20 or DBP >= 10			

The orthostatic vital sign - the change from supine to standing (Supine – Standing) in heart rate, systolic and diastolic blood pressure – will be summarized by scheduled assessment timepoint.

Any vital signs results considered clinically significant by the investigator will be captured as adverse events, hence will show up in AE displays.

9.4.4 Electrocardiogram

Supine 12-lead ECGs will be performed in triplicate, and are collected on screening, days 1, 15, 42, and 182. The following ECG parameters will be listed for each subject: heart rate (beats per minute), PR (msec), QRS (msec), QT (msec), and QTcF (msec).

The average of the triplicate values will be used in the summary, including baseline ECG values. The observed value at each time point and change from baseline at each post-baseline scheduled time point will be summarized. This summary will also include the last values on treatment and on study.

Each ECG is evaluated as 'normal', 'abnormal, not clinically significant' and 'abnormal, clinically significant'; the number and percentage of subjects with at least one of the triplicate values in the categories of 'abnormal, clinically significant' and 'abnormal, not clinically significant' will be provided at baseline and each post-baseline scheduled assessment time point.

Additionally, the number and percentage of subjects with PCS and potentially clinically significant change (PCSC) values will be summarized for such occurrence any time post-baseline. Potentially clinically significant values will be identified for ECG parameters as outlined in the following table. This analysis includes triplicate values individually, and is not based on average value. In addition, the maximum value of QTcF if within any of the PCS criteria will be summarized.

ECG	Units	Criteria for PCS Values (Observed values)		Criteria for PCSC values (Change from Baseline)	
		High	Low	Increase	Decrease
QTcF Interval	msec	>450 but <=480 >480 but <=500 >500	NA	>=30 to 60 >60	NA

9.4.5 Physical Examination

Physical examination is scheduled on screening, day 1 and 42. Only clinically significant abnormalities are captured in the database – for post-baseline observations, these will be reported as adverse events, hence these will be included in AE displays; for pre-baseline observations, these will be reported as medical history, hence these will be included in Medical History displays. The dates of physical examination will be listed to confirm that the examination was done.

9.4.6 Columbia Suicide Severity Rating Scale (C-SSRS)

Suicidality data collected on the C-SSRS is collected during the clinical visits at Screening, Days 1, 3, 8, 12, 15, 21, 28, 35, and 42. 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). The subject's non-suicidal self-injurious behaviors is also assessed separately as part of C-SSRS. The "Baseline/Screening" C-SSRS form will be completed at screening (lifetime history and past 24 months). The "Since Last Visit" C-SSRS form will be completed at all subsequent time points.

The assessments for suicidal ideation is ranked as follows with 5 being the worst:

1. Wish to be dead
2. Non-specific active suicidal thoughts
3. Active suicidal ideation with any methods
4. Active suicidal ideation with some intent
5. Active suicidal ideation with specific plan

The assessments for suicidal behavior is ranked as follows with 5 being the worst:

1. Preparatory acts or behavior
2. Aborted attempt
3. Interrupted attempt
4. Actual attempt (non-fatal)
5. Completed suicide

Suicidal behavior is considered worse than suicidal ideation.

Baseline for each question is defined as the worst of the assessments done before the first dose of study drug, excluding the lifetime version. This will typically include the 'past 24-month 'version

from screening and 'since last visit version' from Day 1, as well as any unscheduled visits done before the first dose of study drug; any Yes will make the baseline value as Yes.

The number and percentage of subjects with at least one response of 'Yes' to any C-SSRS suicidal ideation or suicidal behavior item, as well as for Subject's non-suicidal self-injurious behavior, will be summarized first by visit, then separately for baseline and any time post-baseline.

Summary of shift from baseline in C-SSRS suicidal ideation and suicidal behavior will be presented for the following categories (no suicidal ideation/behavior, suicidal ideation, suicidal behavior) for each scheduled assessment time point. If the answer to all 5 assessments in suicidal ideation and all 5 assessments in suicidal behavior is 'No' then the category for the table is considered as 'No suicidal ideation/behavior'. If any of the assessments in suicidal behavior is Yes, the category is considered as 'Suicidal behavior'. If any of the assessments in suicidal ideation is Yes but all assessments in suicidal behavior is No, the category is considered as 'Suicidal ideation'.

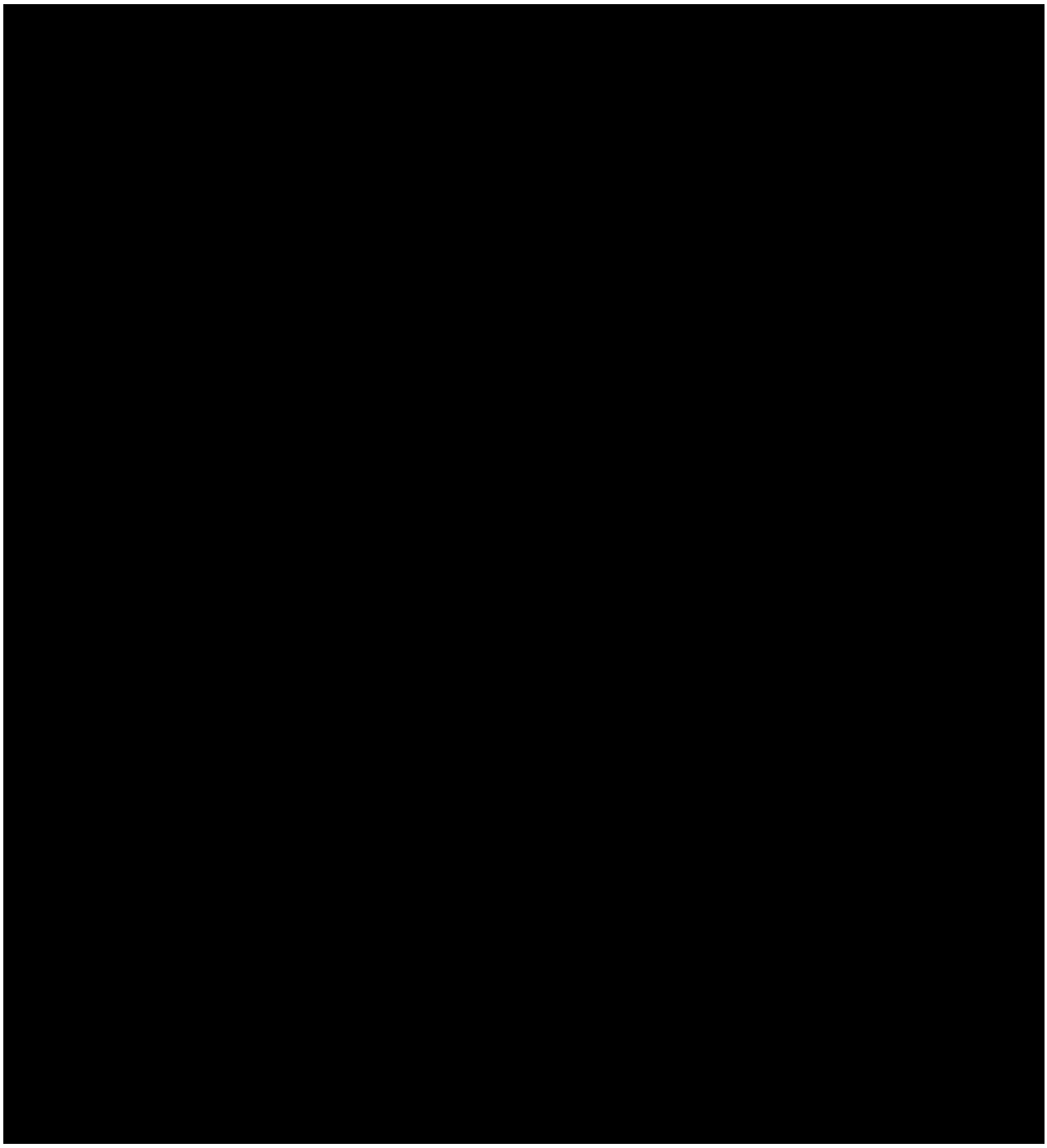
In addition, a summary of shift in suicidal ideation from baseline maximum rank score for any time post-baseline maximum rank score will be presented. Maximum score 0 refers to all No for all assessments in the desired period for all 5 questions on suicidal ideation.

9.4.7 Physician Withdrawal Checklist (PWC)

The PWC-20 will be used to monitor for the presence of potential withdrawal symptoms following discontinuation of SAGE-217

The PWC-20 is made up of a list of 20 symptoms (eg, loss of appetite, nausea-vomiting, diarrhea, anxiety-nervousness, irritability, etc.) that are rated on a scale of 0 (not present) to 3 (severe. The total score is calculated as the sum of 20 responses and ranges from 0 to 60. If more than 2 individual items are missing, the PWC 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 or the maximum possible values for the missing responses, whichever is smaller, to calculate the PWC total score. A higher total score indicates a greater degree withdrawal symptoms.

The observed and change from baseline will be summarized for PWC total score.



10 SUMMARY OF INTERIM AND DMC ANALYSES

Not applicable

11 REFERENCES

Draft clinical study protocol, version 5.0, 25 March 2019, Company: Sage Therapeutics Inc.

[redacted]

12 LIST OF APPENDICES

12.1 Appendix A: Schedule of Assessments

Visits	**Screening Period**	**Double-Blind, Placebo-Controlled Treatment Period**					**Follow-up Period**					**Extended Follow-up**
Visit Days	**D-28 to D-1**	**D1**	**D3 (±1d)**	**D8 (+1d)**	**D12 (±1d)**	**D15 (±1d) and/or EOT[a]**	**D18 (±1d)**	**D21 (±1d)**	**D28 (±3d)**	**D35 (±3d)**	**D42 (±3d) or ET**	**D70, D146, D182** (±7d)
Visit Number	**V1**	**V2**	**V3**	**V4**	**V5**	**V6**	**V7**	**V8**	**V9**	**V10**	**V11**	**V12, V13, V14**
Study Procedure												
Informed Consent	X											
Duplicate Subject Check[b]	X											
Inclusion/Exclusion	X	X										
Serum FSH test[c]	X											
SCID-5-CT	X											
MGH ATRQ	X											
Demographics	X											
Medical/Family History	X											
Subject training[d]	X	X										
Randomization		X										
Physical Examination[e]	X	X									X	
Body Weight/Height	X					X (wt only)					X (wt only)	
Clinical Laboratory Assessments[f]	X	X		X		X		X	X		X	
Drug & Alcohol Screen[g]	X	X	X	X	X	X	X	X	X	X	X	

Visits	**Screening Period**	**Double-Blind, Placebo-Controlled Treatment Period**					**Follow-up Period**					**Extended Follow-up**
Visit Days	**D-28 to D-1**	**D1**	**D3 (±1d)**	**D8 (+1d)**	**D12 (±1d)**	**D15 (±1d) and/or EOT[a]**	**D18 (±1d)**	**D21 (±1d)**	**D28 (±3d)**	**D35 (±3d)**	**D42 (±3d) or ET**	**D70, D146, D182** (±7d)
Visit Number	**V1**	**V2**	**V3**	**V4**	**V5**	**V6**	**V7**	**V8**	**V9**	**V10**	**V11**	**V12, V13, V14**
Study Procedure												
Pregnancy Test[h]	X	X				X			X		X	
Hepatitis & HIV Screen	X											
[redacted]												
Vital Signs[k]	X	X	X	X	X	X	X	X	X	X	X	X
12-Lead ECG[l]	X	X				X					X	X (Day 182 only)
C-SSRS[m]	X	X	X	X	X	X	X	X	X	X	X	X
HAM-D[n, o]		X	X	X	X	X	X	X	X	X	X	X
MADRS	X	X	X	X	X	X	X	X	X	X	X	
HAM-A[o]		X		X		X	X		X		X	
CGI-S	X	X	X	X	X	X	X	X	X	X	X	X
CGI-I			X	X	X	X	X	X	X	X	X	X
SF-36v2	X	X		X		X			X		X	X
PHQ-9		X		X		X		X			X	X
[redacted]												
ISI		X		X		X	X	X	X		X	X
PWC-20		X				X	X	X				

Visits	Screening Period	Double-Blind, Placebo-Controlled Treatment Period					Follow-up Period					Extended Follow-up
Visit Days	D-28 to D-1	D1	D3 (±1d)	D8 (+1d)	D12 (±1d)	D15 (±1d) and/or EOT[a]	D18 (±1d)	D21 (±1d)	D28 (±3d)	D35 (±3d)	D42 (±3d) or ET	D70, D146, D182 (±7d)
Visit Number	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12, V13, V14
Study Procedure												
Sleep diary[p]	X											
[redacted]												
Study Drug Dispensation		X		X								
Study Drug Administration		X (Day 1 through Day 14)										
Study Drug Accountability/Return				X		X					X[r]	
Adverse Events/SAEs[s]	X											
Prior/Concomitant Medications/Procedures[t]	X											

CGI-I = Clinical Global Impression – Improvement; CGI-S – Clinical Global Impression – Severity; [redacted] C-SSRS = Columbia Suicide Severity Rating Scale; D = day; EOT = end of treatment; ET = early termination; ECG = electrocardiogram; [redacted] FSH = follicle stimulating hormone; HAM-A = Hamilton Anxiety Rating Scale; HAM-D = Hamilton Rating Scale for Depression, 17-item; HIV = human immunodeficiency virus; ISI = Insomnia Severity Index; MADRS = Montgomery-Åsberg Depression Rating Scale; MGH ATRQ = Massachusetts General Hospital Antidepressant Treatment Response Questionnaire; PHQ-9 = 9-item Patient Health Questionnaire; PWC-20 = 20-item Physician Withdrawal Checklist; O = Optional; [redacted] SCID-5-CT = Structured Clinical Interview for Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition Clinical Trials Version; SF-36v2 = 36-item Short Form survey version 2; V = visit; wt = weight

[a] Subjects who discontinue treatment early should return to the site for an end of treatment (EOT) visit as soon as possible, preferably the day after treatment is discontinued. Follow-up visits should take place every 7 days after the last dose of treatment for a total of 4 follow-up visits. If at any time after the EOT visit, a subject decides to terminate the study, the subject should return for an early termination (ET) visit. The EOT and ET visits can be on the same day if a subject discontinues study drug and terminates the study on the same day during a clinic visit; in this case, all events scheduled for both visits will be conducted.

[b] Subjects will be asked to authorize that their unique subject identifiers be entered into a registry (www.subjectregistry.com) with the intent of identifying subjects who may meet exclusion criteria for participation in another clinical study.

[c] A serum follicle stimulating hormone test will be conducted at Screening for female subjects that are not surgically sterile to confirm whether a female subject with ≥12 months of spontaneous amenorrhea meets the protocol-defined criteria for being post-menopausal.

[d] Subjects will be trained on use of software applications and devices necessary for the conduct of the study by site personnel.

[e] A full physical examination will be conducted at Screening and abbreviated physical examinations will be conducted thereafter. A full physical examination includes assessment of body systems (eg, head, eye, ear, nose, and throat; heart; lungs; abdomen; and extremities).
[f] Safety laboratory tests will include hematology, serum chemistry, coagulation, and urinalysis.
[g] Urine toxicology for selected drugs of abuse (as per the lab manual) and breath test for alcohol.
[h] Serum pregnancy test at screening and urine pregnancy test thereafter for female subjects that are not surgically sterile and do not meet the protocol-defined criteria for being post-menopausal.
[redacted]
[k] Vital signs include oral temperature (°C), respiratory rate, heart rate, and blood pressure (supine and standing). Heart rate and blood pressure to be collected in supine position at all scheduled time points after the subject has been resting for 5 minutes and then after 1 minute in the standing position. Vital signs may be repeated at the discretion of the Investigator as clinically indicated.
[l] Triplicate ECGs will be collected. [redacted]
[m] The "Baseline/Screening" C-SSRS form will be completed at screening. The "Since Last Visit" C-SSRS form will be completed at any time of day at all subsequent time points.
[n] The HAM-D is to be completed as early during the visit as possible.
[o] The assessment timeframe for HAM-D scales will refer to the past 7 days (1 week) at Screening and "Since Last Visit" for all other visits. The assessment timeframe for HAM-A scale will refer to the past 7 days (1 week) at all visits.
[p] Subjects are instructed to complete the Core Consensus Sleep Diary starting at least 7 days prior to Day 1 and then daily through Day 28.
[redacted]
[r] To be performed at the ET visit only.
[s] Adverse events will be collected starting at the time of informed consent and throughout the duration of the subject's participation in the study.
[t] Prior medications will be collected at Screening and concomitant medications and/or procedures will be collected at each subsequent visit.

12.2 Appendix B: Details of Statistical Methodology

Sample SAS code for Mixed Effects Model for Repeated Measures (MMRM):

```
ods output lsmeans=estimates diffs=diffs;
proc mixed data=&data;
class trtpn avisitn usubjid antidep;
model chg=base trtpn avisitn trtpn*avisitn antidep / ddfm=kr s;
repeated avisitn / subject=usubjid type=un;
**If type=un does not converge, use type=TOEP;
lsmeans trtpn*avisitn /diff=all cl alpha=0.05;
** assuming the Placebo is A, SAGE-217 20 mg is B, SAG-271 30 mg is C
run;
```

Sample SAS code for Generalized Estimating Equation (GEE):

```
proc genmod data=&data;
class usubjid trtpn antidep avisitn;
model aval=base trtpn avisitn trtpn*avisitn antidep /dist=bin link=logit;
repeated subject=usubjid / type=un; * if convergence not met, use type=exch;
lsmeans trtpn*avisitn / diff exp cl;
run;
```

Sample SAS code for Multiple Imputation (MI):

```
** Missing category 1, trtp=A represent the PLACEBO group
proc mi data=in seed=xxxx nimpute=20 round=. . . . 1 1 1 1 1 1 1 1 1 1 1 1 1 out= fcs_reg1;
class trtp strata;
FCS reg;
mnar model (day3 day8 day 12 day15 day18 day21 day28 day42/ modelobs= (trtp='A'));
var strata base day3 day8 day12 day15 day18 day21 day28 day35 day42;
run;
** Missing category 2
```

```
proc mi data=&data seed=xxxx nimpute=20 round=. . . . 1 1 1 1 1 1 1 1 1 1 1 1 1 out=fcs_reg2;
class trtp strata;
fcs nbiter=20 reg (base Day3 Day8  Day12 day15 day18 day21 day28 day35 day42/details);
var trtp strata base Day3 Day8  Day12 day15 day18 day21 day28 day35 day42;
run;
```

12.3 Appendix C: Handling of Missing Dates

Dates missing the day or both the day and month of the year will adhere to the following conventions in order to classify TEAEs and to classify prior and concomitant medications.

In general, listings will present the actual partial or missing values rather than the imputed values that may be used in derivation. In instances where imputed values will be presented, imputed values will be flagged.

12.3.1 Adverse Events

If the AE start date is completely missing, do not impute a date but consider it as TEAE, unless the AE end date is before the initiation of treatment, in which case the AE will be considered prior.

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 + 1 day.
  - 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.

- 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.

- When the year and day are present and the month is missing, treat it as if the day is missing, and only year is present. Follow the imputation rules for "year is known, but the month and day is unknown".
- When the year is missing, but the month and/or day is known, treat this date as missing; do not impute.

12.3.2 Dates in Disease History (Dates of diagnosis, current episode, first episode)

- 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

12.3.3 Prior and Concomitant Medications

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.
- If the year and day are present and the month is missing, then treat it as if the day is also missing. Set the month and day to be December 31.

12.4 Appendix D: List of Displays

Tables

Table Number	Title	Analysis Set
Table 14.1.1.1	Summary of Subject Disposition	All Subjects

Table Number	Title	Analysis Set
Table 14.1.1.2	Summary of Subject Disposition	Modified Full Analysis Set
Table 14.1.2.1	Summary of Major Protocol Deviations	Safety Set
Table 14.1.2.2	Reasons for Exclusion from Analysis Sets	Randomized Set
Table 14.1.3.1.1	Summary of Demographics and Baseline Characteristics	Safety Set
Table 14.1.3.1.2	Summary of Demographics and Baseline Characteristics	Full Analysis Set
Table 14.1.3.1.3	Summary of Demographics and Baseline Characteristics	Modified Full Analysis Set
Table 14.1.3.1.4	Summary of Baseline Subgroups	Safety Set
Table 14.1.3.1.5	Summary of Baseline Subgroups	Full Analysis Set
Table 14.1.3.1.6	Summary of Baseline Subgroups	Modified Full Analysis Set
Table 14.1.3.2.1	Summary of Disease History	Safety Set
Table 14.1.3.3.1	Summary of Medical and Surgical History	Safety Set
Table 14.1.3.3.2	Summary of Medical and Surgical History ongoing at Screening	Safety Set
Table 14.1.3.3.3	Summary of Subject History of Psychiatric Disorder	Safety Set
Table 14.1.3.3.4	Summary of Family History of Psychiatric Disorder	Safety Set
Table 14.1.4.1	Summary of Prior Non-Psychotropic Medications	Safety Set
Table 14.1.4.2	Summary of Concomitant Non-Psychotropic Medications	Safety Set
Table 14.1.4.3	Summary of On-treatment Non-Psychotropic Medications	Safety Set
Table 14.1.4.4	Summary of Post-treatment Non-Psychotropic Medications	Safety Set
Table 14.1.4.5.1	Summary of Prior Psychotropic Medications	Safety Set
Table 14.1.4.5.2	Summary of Prior Psychotropic Medications by ATC Level 4	Safety Set
Table 14.1.4.6.1	Summary of Concomitant Psychotropic Medications	Safety Set
Table 14.1.4.6.2	Summary of Concomitant Psychotropic Medications by ATC Level 4	Safety Set
Table 14.1.4.7	Summary of On-treatment Psychotropic Medications	Safety Set
Table 14.1.4.8	Summary of Post-treatment Psychotropic Medications	Safety Set
Table 14.1.4.9	Summary of Use of Concomitant Antidepressant Medications	Safety Set
Table 14.1.4.10	Summary of Study Drug Exposure	Safety Set
Table 14.1.4.11	Summary of Study Drug Adherence	Full Analysis Set
Table 14.1.4.12	Summary of Study Drug Adherence	Modified Full Analysis Set
Table 14.2.1.1.1	Summary of Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit	Full Analysis Set
Table 14.2.1.1.2	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Total Score, Change from Baseline by Study Visit	Full Analysis Set

Table Number	Title	Analysis Set
Table 14.2.1.1.3	Model-based Sensitivity Analysis on Hamilton Rating Scale for Depression (HAM-D) Total Score, Change from Baseline at Day 15	Full Analysis Set
Table 14.2.1.1.4	Summary of Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit	Per Protocol Set
Table 14.2.1.1.5	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Total Score, Change from Baseline by Study Visit	Per Protocol Set
Table 14.2.1.1.6	Summary of Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit	Modified Full Analysis Set
Table 14.2.1.1.7	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Total Score, Change from Baseline by Study Visit	Modified Full Analysis Set
Table 14.2.1.1.8	Model-based Sensitivity Analysis on Hamilton Rating Scale for Depression (HAM-D) Total Score, Change from Baseline at Day 15	Modified Full Analysis Set
Table 14.2.2.1.1	Summary of Hamilton Rating Scale for Depression (HAM-D) Subscale Scores by Study Visit	Full Analysis Set
Table 14.2.2.1.2	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Core Subscale Score, Change from Baseline by Study Visit	Full Analysis Set
Table 14.2.2.1.3	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Anxiety Subscale Score, Change from Baseline by Study Visit	Full Analysis Set
Table 14.2.2.1.4	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Bech-6 Subscale Score, Change from Baseline by Study Visit	Full Analysis Set
Table 14.2.2.1.5	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Maier Subscale Score, Change from Baseline by Study Visit	Full Analysis Set
Table 14.2.2.1.6	Summary of Hamilton Rating Scale for Depression (HAM-D) Subscale Scores by Study Visit	Modified Full Analysis Set
Table 14.2.2.1.7	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Core Subscale Score, Change from Baseline by Study Visit	Modified Full Analysis Set
Table 14.2.2.1.8	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Anxiety Subscale Score, Change from Baseline by Study Visit	Modified Full Analysis Set
Table 14.2.2.1.9	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Bech-6 Subscale Score, Change from Baseline by Study Visit	Modified Full Analysis Set
Table 14.2.2.1.10	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Maier Subscale Score, Change from Baseline by Study Visit	Modified Full Analysis Set

Table Number	Title	Analysis Set
Table 14.2.2.2.1	Summary of Hamilton Rating Scale for Depression (HAM-D) Individual Item Score by Study Visit	Full Analysis Set
Table 14.2.2.2.2	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Individual Item Score, Change from Baseline by Study Visit	Full Analysis Set
Table 14.2.2.2.3	Summary of Hamilton Rating Scale for Depression (HAM-D) Individual Item Score by Study Visit	Modified Full Analysis Set
Table 14.2.2.2.4	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Individual Item Score, Change from Baseline by Study Visit	Modified Full Analysis Set
Table 14.2.2.3.1	Summary of Hamilton Rating Scale for Depression (HAM-D) Response by Study Visit	Full Analysis Set
Table 14.2.2.3.2.1	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Response by Study Visit	Full Analysis Set
Table 14.2.2.3.2.2	Sensitivity Analysis: Summary of Hamilton Rating Scale for Depression (HAM-D) Response by Study Visit	Full Analysis Set
Table 14.2.2.3.2.3	Summary of Hamilton Rating Scale for Depression (HAM-D) Total Score – Percent Improvement – by Study Visit	Full Analysis Set
Table 14.2.2.3.2.4	Summary of Time to First Hamilton Rating Scale for Depression (HAM-D) Response – Kaplan-Meier Analysis	Full Analysis Set
Table 14.2.2.3.3.1	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Response by Study Visit	Per Protocol Set
Table 14.2.2.3.3.2	Summary of Hamilton Rating Scale for Depression (HAM-D) Response by Study Visit	Modified Full Analysis Set
Table 14.2.2.3.3.3	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Response by Study Visit	Modified Full Analysis Set
Table 14.2.2.3.3.4	Sensitivity Analysis: Summary of Hamilton Rating Scale for Depression (HAM-D) Response by Study Visit	Modified Full Analysis Set
Table 14.2.2.3.3.5	Summary of Hamilton Rating Scale for Depression (HAM-D) Total Score – Percent Improvement – by Study Visit	Modified Full Analysis Set
Table 14.2.2.3.3.6	Summary of Time to First Hamilton Rating Scale for Depression (HAM-D) Response – Kaplan-Meier Analysis	Modified Full Analysis Set
Table 14.2.2.3.4	Summary of Hamilton Rating Scale for Depression (HAM-D) Remission by Study Visit	Full Analysis Set

Table Number	Title	Analysis Set
Table 14.2.2.3.5.1	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Remission by Study Visit	Full Analysis Set
Table 14.2.2.3.5.2	Sensitivity Analysis: Summary of Hamilton Rating Scale for Depression (HAM-D) Remission by Study Visit	Full Analysis Set
Table 14.2.2.3.5.3	Summary of Hamilton Rating Scale for Depression (HAM-D) Total Score in Categories, by Study Visit	Full Analysis Set
Table 14.2.2.3.5.4	Summary of Time to First Hamilton Rating Scale for Depression (HAM-D) Remission – Kaplan-Meier Analysis	Full Analysis Set
Table 14.2.2.3.5.5	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Remission by Study Visit	Per Protocol Set
Table 14.2.2.3.5.6	Summary of Hamilton Rating Scale for Depression (HAM-D) Remission by Study Visit	Modified Full Analysis Set
Table 14.2.2.3.5.7	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Remission by Study Visit	Modified Full Analysis Set
Table 14.2.2.3.5.8	Sensitivity Analysis: Summary of Hamilton Rating Scale for Depression (HAM-D) Remission by Study Visit	Modified Full Analysis Set
Table 14.2.2.3.5.9	Summary of Hamilton Rating Scale for Depression (HAM-D) Total Score in Categories, by Study Visit	Modified Full Analysis Set
Table 14.2.2.3.5.10	Summary of Time to First Hamilton Rating Scale for Depression (HAM-D) Remission – Kaplan-Meier Analysis	Modified Full Analysis Set
Table 14.2.2.4.1	Summary of HAM-D Total Score by Study Visit and Antidepressant Use at Baseline	Full Analysis Set
Table 14.2.2.4.2	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit and Antidepressant Use at Baseline	Full Analysis Set
Table 14.2.2.4.3	Summary of HAM-D Total Score by Study Visit and Age Group	Full Analysis Set
Table 14.2.2.4.4	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit and Age Group	Full Analysis Set
Table 14.2.2.4.5	Summary of HAM-D Total Score by Study Visit and Gender	Full Analysis Set
Table 14.2.2.4.6	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit and Gender	Full Analysis Set
Table 14.2.2.4.7	Summary of HAM-D Total Score by Study Visit and Race Group	Full Analysis Set
Table 14.2.2.4.8	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit and Race Group	Full Analysis Set

Table Number	Title	Analysis Set
Table 14.2.2.4.9	Summary of HAM-D Total Score by Study Visit and BMI Group	Full Analysis Set
Table 14.2.2.4.10	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit and BMI Group	Full Analysis Set
Table 14.2.2.4.11	Summary of HAM-D Total Score by Study Visit and Baseline HAM-D Total Score Category	Full Analysis Set
Table 14.2.2.4.12	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit and Baseline HAM-D Total Score Category	Full Analysis Set
Table 14.2.2.4.13	Summary of HAM-D Total Score by Study Visit and Antidepressant Use at Baseline	Modified Full Analysis Set
Table 14.2.2.4.14	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit and Antidepressant Use at Baseline	Modified Full Analysis Set
Table 14.2.2.4.15	Summary of HAM-D Total Score by Study Visit and Age Group	Modified Full Analysis Set
Table 14.2.2.4.16	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit and Age Group	Modified Full Analysis Set
Table 14.2.2.4.17	Summary of HAM-D Total Score by Study Visit and Gender	Modified Full Analysis Set
Table 14.2.2.4.18	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit and Gender	Modified Full Analysis Set
Table 14.2.2.4.19	Summary of HAM-D Total Score by Study Visit and Race Group	Modified Full Analysis Set
Table 14.2.2.4.20	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit and Race Group	Modified Full Analysis Set
Table 14.2.2.4.21	Summary of HAM-D Total Score by Study Visit and BMI Group	Modified Full Analysis Set
Table 14.2.2.4.22	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit and BMI Group	Modified Full Analysis Set
Table 14.2.2.4.23	Summary of HAM-D Total Score by Study Visit and Baseline HAM-D Total Score Category	Modified Full Analysis Set
Table 14.2.2.4.24	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit and Baseline HAM-D Total Score Category	Modified Full Analysis Set
Table 14.2.2.5.1	Summary of Clinical Global Impression – Improvement (CGI-I) by Study Visit	Full Analysis Set
Table 14.2.2.5.2	Summary of Clinical Global Impression – Improvement (CGI-I) Response by Study Visit	Full Analysis Set
Table 14.2.2.5.3	Model-based Results on Clinical Global Impression – Improvement (CGI-I) Response by Study Visit	Full Analysis Set
Table 14.2.2.5.4	Summary of Clinical Global Impression – Severity (CGI-S) by Study Visit	Full Analysis Set

Table Number	Title	Analysis Set
Table 14.2.2.5.5	Model-based Results on Clinical Global Impression – Severity (CGI-S), Change from Baseline by Study Visit	Full Analysis Set
Table 14.2.2.5.6	Summary of Clinical Global Impression – Improvement (CGI-I) by Study Visit	Modified Full Analysis Set
Table 14.2.2.5.7	Summary of Clinical Global Impression – Improvement (CGI-I) Response by Study Visit	Modified Full Analysis Set
Table 14.2.2.5.8	Model-based Results on Clinical Global Impression – Improvement (CGI-I) Response by Study Visit	Modified Full Analysis Set
Table 14.2.2.5.9	Summary of Clinical Global Impression – Severity (CGI-S) by Study Visit	Modified Full Analysis Set
Table 14.2.2.5.10	Model-based Results on Clinical Global Impression – Severity (CGI-S), Change from Baseline by Study Visit	Modified Full Analysis Set
Table 14.2.2.6.1	Summary of Hamilton Rating Scale for Anxiety (HAM-A) Total Score by Visit	Full Analysis Set
Table 14.2.2.6.2	Model-based Results on Hamilton Rating Scale for Depression (HAM-A) Total Score, Change from Baseline by Study Visit	Full Analysis Set
Table 14.2.2.6.3	Summary of Hamilton Rating Scale for Anxiety (HAM-A) Individual Item Score by Study Visit	Full Analysis Set
Table 14.2.2.6.4	Summary of Hamilton Rating Scale for Anxiety (HAM-A) Total Score by Visit	Modified Full Analysis Set
Table 14.2.2.6.5	Model-based Results on Hamilton Rating Scale for Depression (HAM-A) Total Score, Change from Baseline by Study Visit	Modified Full Analysis Set
Table 14.2.2.6.6	Summary of Hamilton Rating Scale for Anxiety (HAM-A) Individual Item Score by Study Visit	Modified Full Analysis Set
Table 14.2.2.7.1	Summary of Montgomery-Asberg Depression Rating Scale (MADRS) Total Score by Study Visit	Full Analysis Set
Table 14.2.2.7.2	Model-based Results on Montgomery-Asberg Depression Rating Scale (MADRS) Total Score, Change from Baseline by Study Visit	Full Analysis Set
Table 14.2.2.7.3	Summary of Montgomery-Asberg Depression Rating Scale (MADRS) Individual Item Score by Study Visit	Full Analysis Set
Table 14.2.2.7.3.1	Model-based Results on Montgomery-Asberg Depression Rating Scale (MADRS) Individual Item Score, Change from Baseline	Full Analysis Set
Table 14.2.2.7.4	Summary of Montgomery-Asberg Depression Rating Scale (MADRS) Total Score by Study Visit	Per Protocol Set
Table 14.2.2.7.5.1	Summary of Montgomery-Asberg Depression Rating Scale (MADRS) Total Score by Study Visit	Modified Full Analysis Set

Table Number	Title	Analysis Set
Table 14.2.2.7.5.2	Model-based Results on Montgomery-Asberg Depression Rating Scale (MADRS) Total Score, Change from Baseline by Study Visit	Modified Full Analysis Set
Table 14.2.2.7.5.3	Summary of Montgomery-Asberg Depression Rating Scale (MADRS) Individual Item Score by Study Visit	Modified Full Analysis Set
Table 14.2.2.7.5.4	Model-based Results on Montgomery-Asberg Depression Rating Scale (MADRS) Individual Item Score, Change from Baseline	Modified Full Analysis Set
Table 14.2.2.7.6	Summary of Montgomery-Asberg Depression Rating Scale (MADRS) Response by Study Visit	Full Analysis Set
Table 14.2.2.7.7	Model-based Results on Montgomery-Asberg Depression Rating Scale (MADRS) Response by Study Visit	Full Analysis Set
Table 14.2.2.7.8	Summary of Montgomery-Asberg Depression Rating Scale (MADRS) Remission by Study Visit	Full Analysis Set
Table 14.2.2.7.9	Model-based Results on Montgomery-Asberg Depression Rating Scale (MADRS) Remission by Study Visit	Full Analysis Set
Table 14.2.2.7.10	Summary of Montgomery-Asberg Depression Rating Scale (MADRS) Response by Study Visit	Modified Full Analysis Set
Table 14.2.2.7.11	Model-based Results on Montgomery-Asberg Depression Rating Scale (MADRS) Response by Study Visit	Modified Full Analysis Set
Table 14.2.2.7.12	Summary of Montgomery-Asberg Depression Rating Scale (MADRS) Remission by Study Visit	Modified Full Analysis Set
Table 14.2.2.7.13	Model-based Results on Montgomery-Asberg Depression Rating Scale (MADRS) Remission by Study Visit	Modified Full Analysis Set
Table 14.2.2.8.1	Summary of Insomnia Severity Index (ISI) Total Score by Study Visit	Full Analysis Set
Table 14.2.2.8.2	Model-based Results on Insomnia Severity Index (ISI) Total Score, Change from Baseline by Study Visit	Full Analysis Set
Table 14.2.2.8.3	Summary of Insomnia Severity Index (ISI) Total Score by Study Visit	Modified Full Analysis Set
Table 14.2.2.8.4	Model-based Results on Insomnia Severity Index (ISI) Total Score, Change from Baseline by Study Visit	Modified Full Analysis Set
Table 14.2.2.9.1	Summary of Core Consensus Sleep Diary by Study Visit	Full Analysis Set
Table 14.2.2.9.2	Model-based Results on Subjective Sleep Latency (sSL) , Change from Baseline by Study Visit	Full Analysis Set
Table 14.2.2.9.3	Model-based Results on Subjective Wake after Sleep Onset (sWASO) , Change from Baseline by Study Visit	Full Analysis Set

Table Number	Title	Analysis Set
Table 14.2.2.9.4	Model-based Results on Subjective Number of Awakenings (sNAW) , Change from Baseline by Study Visit	Full Analysis Set
Table 14.2.2.9.5	Model-based Results on Subjective Total Sleep Time (sTST) , Change from Baseline by Study Visit	Full Analysis Set
Table 14.2.2.9.6	Model-based Results on Subjective Sleep Quality (sSQ) Response by Study Visit	Full Analysis Set
Table 14.2.2.9.7	Summary of Core Consensus Sleep Diary by Study Visit	Modified Full Analysis Set
Table 14.2.2.9.8	Model-based Results on Subjective Sleep Latency (sSL) , Change from Baseline by Study Visit	Modified Full Analysis Set
Table 14.2.2.9.9	Model-based Results on Subjective Wake after Sleep Onset (sWASO) , Change from Baseline by Study Visit	Modified Full Analysis Set
Table 14.2.2.9.10	Model-based Results on Subjective Number of Awakenings (sNAW) , Change from Baseline by Study Visit	Modified Full Analysis Set
Table 14.2.2.9.11	Model-based Results on Subjective Total Sleep Time (sTST) , Change from Baseline by Study Visit	Modified Full Analysis Set
Table 14.2.2.9.12	Model-based Results on Subjective Sleep Quality (sSQ) Response by Study Visit	Modified Full Analysis Set
Table 14.2.2.10.1	Summary of Short Form-36 Version 2 (SF-36v2) Domain/Component Score by Study Visit	Full Analysis Set
Table 14.2.2.10.2	Model-based Results on Short Form-36 Version 2 (SF-36v2) Physical Functioning Domain Score, Change from Baseline by Study Visit	Full Analysis Set
Table 14.2.2.10.3	Model-based Results on Short Form-36 Version 2 (SF-36v2) Role-physical Domain Score, Change from Baseline by Study Visit	Full Analysis Set
Table 14.2.2.10.4	Model-based Results on Short Form-36 Version 2 (SF-36v2) Bodily Pain Domain Score, Change from Baseline by Study Visit	Full Analysis Set
Table 14.2.2.10.5	Model-based Results on Short Form-36 Version 2 (SF-36v2) General Health Domain Score, Change from Baseline by Study Visit	Full Analysis Set
Table 14.2.2.10.6	Model-based Results on Short Form-36 Version 2 (SF-36v2) Vitality Domain Score, Change from Baseline by Study Visit	Full Analysis Set
Table 14.2.2.10.7	Model-based Results on Short Form-36 Version 2 (SF-36v2) Social Functioning Domain Score, Change from Baseline by Study Visit	Full Analysis Set
Table 14.2.2.10.8	Model-based Results on Short Form-36 Version 2 (SF-36v2) Role-emotional Domain Score, Change from Baseline by Study Visit	Full Analysis Set
Table 14.2.2.10.9	Model-based Results on Short Form-36 Version 2 (SF-36v2) Mental Health Domain Score, Change from Baseline by Study Visit	Full Analysis Set

Table Number	Title	Analysis Set
Table 14.2.2.10.10	Model-based Results on Short Form-36 Version 2 (SF-36v2) Physical Component Summary Score, Change from Baseline by Study Visit	Full Analysis Set
Table 14.2.2.10.11	Model-based Results on Short Form-36 Version 2 (SF-36v2) Mental Component Summary Score, Change from Baseline by Study Visit	Full Analysis Set
Table 14.2.2.10.12	Model-based Results on Short Form-36 Version 2 (SF-36v2) Utility Score (Release 2), Change from Baseline by Study Visit	Full Analysis Set
Table 14.2.2.10.13	Summary of Short Form-36 Version 2 (SF-36v2) Domain/Component Score by Study Visit	Modified Full Analysis Set
Table 14.2.2.10.14	Model-based Results on Short Form-36 Version 2 (SF-36v2) Physical Functioning Domain Score, Change from Baseline by Study Visit	Modified Full Analysis Set
Table 14.2.2.10.15	Model-based Results on Short Form-36 Version 2 (SF-36v2) Role-physical Domain Score, Change from Baseline by Study Visit	Modified Full Analysis Set
Table 14.2.2.10.16	Model-based Results on Short Form-36 Version 2 (SF-36v2) Bodily Pain Domain Score, Change from Baseline by Study Visit	Modified Full Analysis Set
Table 14.2.2.10.17	Model-based Results on Short Form-36 Version 2 (SF-36v2) General Health Domain Score, Change from Baseline by Study Visit	Modified Full Analysis Set
Table 14.2.2.10.18	Model-based Results on Short Form-36 Version 2 (SF-36v2) Vitality Domain Score, Change from Baseline by Study Visit	Modified Full Analysis Set
Table 14.2.2.10.19	Model-based Results on Short Form-36 Version 2 (SF-36v2) Social Functioning Domain Score, Change from Baseline by Study Visit	Modified Full Analysis Set
Table 14.2.2.10.20	Model-based Results on Short Form-36 Version 2 (SF-36v2) Role-emotional Domain Score, Change from Baseline by Study Visit	Modified Full Analysis Set
Table 14.2.2.10.21	Model-based Results on Short Form-36 Version 2 (SF-36v2) Mental Health Domain Score, Change from Baseline by Study Visit	Modified Full Analysis Set
Table 14.2.2.10.22	Model-based Results on Short Form-36 Version 2 (SF-36v2) Physical Component Summary Score, Change from Baseline by Study Visit	Modified Full Analysis Set
Table 14.2.2.10.23	Model-based Results on Short Form-36 Version 2 (SF-36v2) Mental Component Summary Score, Change from Baseline by Study Visit	Modified Full Analysis Set
Table 14.2.2.10.24	Model-based Results on Short Form-36 Version 2 (SF-36v2) Utility Score (Release 2), Change from Baseline by Study Visit	Modified Full Analysis Set
Table 14.2.2.11.1	Summary of Patient Health Questionnaire (PHQ-9) Total Score by Study Visit	Full Analysis Set
Table 14.2.2.11.2	Model-based Results on Patient Health Questionnaire (PHQ-9) Total Score, Change from Baseline by Study Visit	Full Analysis Set

Table Number	Title	Analysis Set
Table 14.2.2.11.3	Summary of Patient Health Questionnaire (PHQ-9) Total Score by Study Visit	Modified Full Analysis Set
Table 14.2.2.11.4	Model-based Results on Patient Health Questionnaire (PHQ-9) Total Score, Change from Baseline by Study Visit	Modified Full Analysis Set
Table 14.3.1.1	Overview of Treatment-Emergent Adverse Events	Safety Set
Table 14.3.1.2	Summary of Treatment-Emergent Adverse Events by System Organ Class and Preferred Term	Safety Set
Table 14.3.1.3	Summary of Double-Blind Period Treatment-Emergent Adverse Events by System Organ Class and Preferred Term	Safety Set
Table 14.3.1.4	Summary of Treatment Period Treatment-Emergent Adverse Events by System Organ Class and Preferred Term	Safety Set
Table 14.3.1.5	Summary of Follow-up Period Treatment-Emergent Adverse Events by System Organ Class and Preferred Term	Safety Set
Table 14.3.1.6	Summary of Extended Follow-up Period Treatment-Emergent Adverse Events by System Organ Class and Preferred Term	Safety Set
Table 14.3.1.7	Summary of Double-Blind Period Treatment-Emergent Adverse Events by System Organ Class, Preferred Term, and Use of Antidepressant at Baseline	Safety Set
Table 14.3.1.8	Summary of Double-Blind Period Treatment-Emergent Adverse Events by System Organ Class, Preferred Term, and Age Group	Safety Set
Table 14.3.1.9	Summary of Double-Blind Period Treatment-Emergent Adverse Events by System Organ Class, Preferred Term, and Gender	Safety Set
Table 14.3.1.10	Summary of Double-Blind Period Treatment-Emergent Adverse Events by System Organ Class, Preferred Term, and Race Group	Safety Set
Table 14.3.1.10.1	Summary of Double-Blind Period Treatment-Emergent Adverse Events by System Organ Class, Preferred Term, and BMI Group	Safety Set
Table 14.3.1.10.2	Summary of Double-Blind Period Treatment-Emergent Adverse Events by System Organ Class, Preferred Term, and Baseline HAM-D Total Score	Safety Set
Table 14.3.1.11	Summary of Double-Blind Period Treatment-Emergent Adverse Events by System Organ Class, Preferred Term, and Maximum Severity	Safety Set
Table 14.3.1.12	Summary of Double-Blind Period Treatment-Emergent Adverse Events by System Organ Class, Preferred Term, and Relationship to Study Drug	Safety Set
Table 14.3.1.13	Summary of Treatment Period Treatment-Emergent Adverse Events Leading to Discontinuation of Study Drug by System Organ Class and Preferred Term	Safety Set

Table Number	Title	Analysis Set
Table 14.3.1.14	Summary of Treatment-Emergent Adverse Events Leading to Withdrawal from the Study by System Organ Class and Preferred Term	Safety Set
Table 14.3.1.15	Summary of Double-Blind Period Treatment-Emergent Adverse Events Leading to Withdrawal from the Study by System Organ Class and Preferred Term	Safety Set
Table 14.3.1.16	Summary of Treatment-Emergent Serious Adverse Events by System Organ Class and Preferred Term	Safety Set
Table 14.3.1.17	Summary of Double-Blind Period Treatment-Emergent Serious Adverse Events by System Organ Class and Preferred Term	Safety Set
Table 14.3.1.18	Summary of Treatment Period Treatment-Emergent Serious Adverse Events by System Organ Class and Preferred Term	Safety Set
Table 14.3.1.19	Summary of Follow-up Period Treatment-Emergent Serious Adverse Events by System Organ Class and Preferred Term	Safety Set
Table 14.3.1.20	Summary of Extended Follow-up Period Treatment-Emergent Serious Adverse Events by System Organ Class and Preferred Term	Safety Set
Table 14.3.1.21	Summary of Most Common (>5%) Double-Blind Period Treatment-Emergent Adverse Events by Preferred Term	Safety Set
Table 14.3.4.1.1.1	Summary of Serum Chemistry Results by Study Visit	Safety Set
Table 14.3.4.1.1.2	Summary of Shift in Serum Chemistry Results during the Study	Safety Set
Table 14.3.4.1.1.3	Summary of Potentially Clinically Significant Liver Function Tests at Any Time Post-Baseline	Safety Set
Table 14.3.4.1.1.4	Summary of Potentially Clinically Significant Serum Chemistry (excluding LFT) at Any Time Post-Baseline	Safety Set
Table 14.3.4.1.2.1	Summary of Hematology Results by Study Visit	Safety Set
Table 14.3.4.1.2.2	Summary of Shift in Hematology Results during the Study	Safety Set
Table 14.3.4.1.2.3	Summary of Potentially Clinically Significant Hematology Results Any Time Post-Baseline	Safety Set
Table 14.3.4.1.3.1	Summary of Urinalysis Results by Study Visit – Quantitative Results	Safety Set
Table 14.3.4.1.3.2	Summary of Urinalysis by Study Visit – Qualitative Results	Safety Set
Table 14.3.4.1.4.1	Summary of Coagulation Results by Study Visit	Safety Set
Table 14.3.4.2.1	Summary of Vital Signs by Study Visit	Safety Set
Table 14.3.4.2.2	Summary of Potentially Clinically Significant Vital Signs Any Time Post-Baseline	Safety Set
Table 14.3.4.2.3	Summary of Orthostatic Changes (Supine minus Standing in Vital Signs by Study Visit	Safety Set

Table Number	Title	Analysis Set
Table 14.3.4.3.1	Summary of ECG data by Study Visit	Safety Set
Table 14.3.4.3.2	Summary of Abnormal ECG by Study Visit	Safety Set
Table 14.3.4.3.3	Summary of Potentially Clinically Significant QTcF in ECG Data Any Time Post-Baseline	Safety Set
Table 14.3.4.4.1	Summary of Columbia Suicide Rating Scale (C-SSRS) Suicidal Ideation and Suicidal Behavior by Study Visit and Treatment Group	Safety Set
Table 14.3.4.4.2	Summary of Columbia-Suicide Severity Rating Scale (C-SSRS) on Suicidal Ideation and Suicidal Behavior Data	Safety Set
Table 14.3.4.4.3.1	Summary of Shift from Baseline in Columbia-Suicide Severity Rating Scale (C-SSRS) Suicidal Ideation and Suicidal Behavior by Study Visit	Safety Set
Table 14.3.4.4.3.2	Summary of Shift from Baseline in C0SSRS Maximum Severity Score in Suicidal Ideation at Any Time Post-Baseline	Safety Set
Table 14.3.4.4.4	Summary of Physical Withdrawal Checklist (PWC) Total Scores by Study Visit	Safety Set

[redacted]

Figures

Figure Number	Title	Analysis Set
Figure 14.2.2.1.2.1	Line Plot of LS Mean (±SE) Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score Over Time by Treatment Group	Full Analysis Set

Figure Number	Title	Analysis Set
Figure 14.2.2.1.4.1	Line Plot of LS Mean (±SE) Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Core Subscale Score Over Time by Treatment Group	Full Analysis Set
Figure 14.2.2.1.5.1	Line Plot of LS Mean (±SE) Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Anxiety Subscale Score Over Time by Treatment Group	Full Analysis Set
Figure 14.2.2.1.6.1	Line Plot of LS Mean (±SE) Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Bech-6 Subscale Score Over Time by Treatment Group	Full Analysis Set
Figure 14.2.2.1.7.1	Line Plot of LS Mean (±SE) Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Maier Subscale Score Over Time by Treatment Group	Full Analysis Set
Figure 14.2.2.1.9.1	Line Plot of LS Mean (±SE) Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score Over Time by Treatment Group	Per Protocol Set
Figure 14.2.2.1.9.2	Line Plot of LS Mean (±SE) Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score Over Time by Treatment Group	Modified Full Analysis Set
Figure 14.2.2.1.9.3	Line Plot of LS Mean (±SE) Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Core Subscale Score Over Time by Treatment Group	Modified Full Analysis Set
Figure 14.2.2.1.9.4	Line Plot of LS Mean (±SE) Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Anxiety Subscale Score Over Time by Treatment Group	Modified Full Analysis Set
Figure 14.2.2.1.9.5	Line Plot of LS Mean (±SE) Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Bech-6 Subscale Score Over Time by Treatment Group	Modified Full Analysis Set
Figure 14.2.2.1.9.6	Line Plot of LS Mean (±SE) Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Maier Subscale Score Over Time by Treatment Group	Modified Full Analysis Set
Figure 14.2.2.1.9.7	Line Plot of LS Mean (±SE) Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score Over Time by Treatment Group and Antidepressant Use at Baseline	Modified Full Analysis Set
Figure 14.2.2.1.9.8	Line Plot of LS Mean (±SE) Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score Over Time by Treatment Group and Age Group	Modified Full Analysis Set
Figure 14.2.2.1.9.9	Line Plot of LS Mean (±SE) Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score Over Time by Treatment Group and Sex	Modified Full Analysis Set
Figure 14.2.2.1.9.10	Line Plot of LS Mean (±SE) Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score Over Time by Treatment Group and Race Group	Modified Full Analysis Set
Figure 14.2.2.1.9.11	Line Plot of LS Mean (±SE) Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score Over Time by Treatment Group and BMI Group	Modified Full Analysis Set

Figure Number	Title	Analysis Set
Figure 14.2.2.1.9.12	Line Plot of LS Mean (±SE) Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score Over Time by Treatment Group and Baseline HAM-D Total Score	Modified Full Analysis Set
Figure 14.2.2.1.10.1	Bar Chart of Hamilton Rating Scale for Depression (HAM-D) Total Score, Response over Time	Full Analysis Set
Figure 14.2.2.1.10.2	Histogram of Hamilton Rating Scale for Depression (HAM-D) Percentage Improvement over Time	Full Analysis Set
Figure 14.2.2.1.10.3	Bar Chart of Hamilton Rating Scale for Depression (HAM-D) Total Score, Remission over Time by Treatment Group	Full Analysis Set
Figure 14.2.2.1.10.3.1	Histogram of Hamilton Rating Scale for Depression (HAM-D) Total Score over Time	Full Analysis Set
Figure 14.2.2.1.10.4	Bar Chart of Hamilton Rating Scale for Depression (HAM-D) Total Score, Response over Time	Modified Full Analysis Set
Figure 14.2.2.1.10.5	Histogram of Hamilton Rating Scale for Depression (HAM-D) Percentage Improvement over Time	Modified Full Analysis Set
Figure 14.2.2.1.10.6	Bar Chart of Hamilton Rating Scale for Depression (HAM-D) Total Score, Remission over Time by Treatment Group	Modified Full Analysis Set
Figure 14.2.2.1.10.6.1	Histogram of Hamilton Rating Scale for Depression (HAM-D) Total Score over Time	Modified Full Analysis Set
Figure 14.2.2.1.11.1	Forest Plot of Change from Baseline at Day 15 in Hamilton Rating Scale for Depression (HAM-D) Total Score for Subgroups	Full Analysis Set
Figure 14.2.2.1.11.2	Forest Plot of Change from Baseline at Day 15 in Hamilton Rating Scale for Depression (HAM-D) Total Score for Subscales of HAM-D	Full Analysis Set
Figure 14.2.2.1.11.3	Forest Plot of Change from Baseline at Day 15 in Hamilton Rating Scale for Depression (HAM-D) Total Score for Individual Items of HAM-D Scale	Full Analysis Set
Figure 14.2.2.1.11.4	Forest Plot of Change from Baseline at Day 15 in Hamilton Rating Scale for Depression (HAM-D) Total Score for Subgroups	Modified Full Analysis Set
Figure 14.2.2.1.11.5	Forest Plot of Change from Baseline at Day 15 in Hamilton Rating Scale for Depression (HAM-D) Total Score for Subscales of HAM-D	Modified Full Analysis Set
Figure 14.2.2.1.11.6	Forest Plot of Change from Baseline at Day 15 in Hamilton Rating Scale for Depression (HAM-D) Total Score for Individual Items of HAM-D Scale	Modified Full Analysis Set
Figure 14.2.2.5.2.1	Bar Chart of CGI-I Response over Time by Treatment Group	Full Analysis Set

Figure Number	Title	Analysis Set
Figure 14.2.2.5.5.1	Line Plot of LS Mean (±SE) Change from Baseline in CGI-S over Time by Treatment Group	Full Analysis Set
Figure 14.2.2.5.7.1	Bar Chart of CGI-I Response over Time by Treatment Group	Modified Full Analysis Set
Figure 14.2.2.5.10.1	Line Plot of LS Mean (±SE) Change from Baseline in CGI-S over Time by Treatment Group	Modified Full Analysis Set
Figure 14.2.2.6.2.1	Line Plot of LS Mean (±SE) Change from Baseline in Hamilton Rating Scale for Anxiety (HAM-A) Total Score Over Time by Treatment Group	Full Analysis Set
Figure 14.2.2.6.2.2	Line Plot of LS Mean (±SE) Change from Baseline in Hamilton Rating Scale for Anxiety (HAM-A) Total Score Over Time by Treatment Group	Modified Full Analysis Set
Figure 14.2.2.7.2.1	Line Plot of LS Mean (±SE) Change from Baseline in Montgomery-Asberg Depression Rating Scale (MADRS) Total Score Over Time by Treatment Group	Full Analysis Set
Figure 14.2.2.7.2.2	Line Plot of LS Mean (±SE) Change from Baseline in Montgomery-Asberg Depression Rating Scale (MADRS) Total Score Over Time by Treatment Group	Modified Full Analysis Set
Figure 14.2.2.8.2.1	Line Plot of LS Mean (±SE) Change from Baseline in Insomnia Severity Index (ISI) Total Score Over Time by Treatment Group	Full Analysis Set
Figure 14.2.2.8.2.2	Line Plot of LS Mean (±SE) Change from Baseline in Insomnia Severity Index (ISI) Total Score Over Time by Treatment Group	Modified Full Analysis Set
Figure 14.2.2.9.2.1	Line Plot of LS Mean (±SE) Change from Baseline in Subjective Sleep Latency (sSL) Total Score Over Time by Treatment Group	Full Analysis Set
Figure 14.2.2.9.3.1	Line Plot of LS Mean (±SE) Change from Baseline in Subjective Total Sleep Time (sTST) Total Score Over Time by Treatment Group	Full Analysis Set
Figure 14.2.2.9.4.1	Line Plot of LS Mean (±SE) Change from Baseline in Subjective Wake after Sleep Onset (sWASO) Total Score Over Time by Treatment Group	Full Analysis Set
Figure 14.2.2.9.5.1	Line Plot of LS Mean (±SE) Change from Baseline in Subjective Number of Awakenings (sNAW) Total Score Over Time by Treatment Group	Full Analysis Set
Figure 14.2.2.9.6.1	Line Plot of LS Mean (±SE) Change from Baseline in Subjective Sleep Latency (sSL) Total Score Over Time by Treatment Group	Modified Full Analysis Set
Figure 14.2.2.9.7.1	Line Plot of LS Mean (±SE) Change from Baseline in Subjective Total Sleep Time (sTST) Total Score Over Time by Treatment Group	Modified Full Analysis Set
Figure 14.2.2.9.8.1	Line Plot of LS Mean (±SE) Change from Baseline in Subjective Wake after Sleep Onset (sWASO) Total Score Over Time by Treatment Group	Modified Full Analysis Set
Figure 14.2.2.9.9.1	Line Plot of LS Mean (±SE) Change from Baseline in Subjective Number of Awakenings (sNAW) Over Time by Treatment Group	Modified Full Analysis Set

Figure Number	Title	Analysis Set
Figure 14.2.2.10.2.1	Bar Chart of LS Mean (±SE) Change from Baseline in Short Form-36 Version 2 (SF-36v2) at Day 15 by Domain/Component and Treatment Group	Full Analysis Set
Figure 14.2.2.10.2.2	Bar Chart of LS Mean (±SE) Change from Baseline in Short Form-36 Version 2 (SF-36v2) at Day 15 by Domain/Component and Treatment Group	Modified Full Analysis Set
Figure 14.2.2.11.2	Model-based LS Mean (±SE) Change from Baseline in Patient Health Questionnaire (PHQ-9) Total Score Over Time by Treatment Group	Full Analysis Set
Figure 14.2.2.11.3	Model-based LS Mean (±SE) Change from Baseline in Patient Health Questionnaire (PHQ-9) Total Score Over Time by Treatment Group	Modified Full Analysis Set

Listings

Listing Number	Title	Analysis Set
Listing 16.2.1.1	Subject Disposition	All Subjects
Listing 16.2.1.2	Premature Discontinuation from Study Drug and/or Withdrawal from the Study	Safety Set
Listing 16.2.2.1	Protocol Deviations	Safety Set
Listing 16.2.2.2	Inappropriate Study Drug Consumption	Safety Set
Listing 16.2.3.1	Inclusion/Exclusion Criteria Violations	Randomized Set
Listing 16.2.4.1	Demographics	Safety Set
Listing 16.2.4.2	Baseline Characteristics	Safety Set
Listing 16.2.4.2.1	Incorrect Stratification related to Antidepressant Use at Baseline	Safety Set
Listing 16.2.4.2.2	Diagnostic Laboratory Results	Safety Set
Listing 16.2.4.3.1	Medical and Surgical History	Safety Set
Listing 16.2.4.3.2	Disease History	Safety Set
Listing 16.2.4.3.3	Subject History of Psychiatric Disorder	Safety Set
Listing 16.2.4.3.4	Family History of Psychiatric Disorder	Safety Set
Listing 16.2.4.4.1	Prior and Concomitant Psychotropic Medications	Safety Set
Listing 16.2.4.4.2	Prior and Concomitant Non-Psychotropic Medications	Safety Set
Listing 16.2.4.4.3	Listing of Use of Antidepressant during the Study	Safety Set
Listing 16.2.4.4.4	Listing of Use of Atypical Antipsychotics during the Study	Safety Set
Listing 16.2.4.4.5	Concomitant Procedures	Safety Set
Listing 16.2.5.1	Study Drug Administration	Safety Set
Listing 16.2.5.2	Study Drug Exposure	Safety Set

Listing Number	Title	Analysis Set
Listing 16.2.5.3	Study Drug Adherence	Safety Set
Listing 16.2.6.1	Hamilton Rating Scale for Depression (HAM-D)	Full Analysis Set
Listing 16.2.6.2.1	Clinical Global Impression (CGI) – Improvement	Full Analysis Set
Listing 16.2.6.2.2	Clinical Global Impression (CGI) – Severity	Full Analysis Set
Listing 16.2.6.2.3	Hamilton Anxiety Rating Scale (HAM-A)	Full Analysis Set
Listing 16.2.6.2.4	Montgomery-Asberg Depression Rating Scale (MADRS)	Full Analysis Set
Listing 16.2.6.2.5	Insomnia Severity Index (ISI)	Full Analysis Set
Listing 16.2.6.2.6	Core Consensus Sleep Diary	Full Analysis Set
Listing 16.2.6.2.7.1	Short-Form 36 (SF-36) Individual Item Scores	Full Analysis Set
Listing 16.2.6.2.7.2	Short-Form 36 (SF-36) Domain/Component Scores	Full Analysis Set
Listing 16.2.6.2.8	Patient Health Questionnaire (PHQ-9)	Full Analysis Set
Listing 16.2.7.1	Treatment-Emergent Adverse Events	Safety Set
Listing 16.2.7.2	Non-fatal Serious Adverse Events	Safety Set
Listing 16.2.7.3	Adverse Events Leading to Death	Safety Set
Listing 16.2.7.4	Adverse Events with Onset Prior to First Dose of Study Drug	Safety Set
Listing 16.2.7.5	Adverse Events leading to Premature Discontinuation of Study Drug	Safety Set
Listing 16.2.7.6	Adverse Events leading to Premature Withdrawal from the Study	Safety Set
Listing 16.2.8.1.1	Serum Chemistry Data (excluding Liver Function Tests)	Safety Set
Listing 16.2.8.1.1.1	Potentially Clinically Significant Serum Chemistry Values (excluding Liver Function Tests)	Safety Set
Listing 16.2.8.1.2	Serum Chemistry Data, Liver Function Tests	Safety Set
Listing 16.2.8.1.3	Hematology Data	Safety Set
Listing 16.2.8.1.4	Urinalysis Data	Safety Set
Listing 16.2.8.1.5	Coagulation Data	Safety Set
Listing 16.2.8.1.6	Pregnancy Test	Safety Set
Listing 16.2.8.2	Vital Signs	Safety Set
Listing 16.2.8.3	Electrocardiogram (ECG) Data	Safety Set
Listing 16.2.8.4	Physical Examination	Safety Set
Listing 16.2.8.5.4.1	Columbia-Suicide Severity Rating Scale (C-SSRS) – Suicidal Ideation	Safety Set
Listing 16.2.8.5.4.2	Columbia-Suicide Severity Rating Scale (C-SSRS) – Suicidal Behavior	Safety Set
Listing 16.2.8.6	Physician Withdrawal Checklist (PWC)	Safety Set

Listing Number	Title	Analysis Set
[redacted]		