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: NCT04442490

Document Date: SAP Version 1.0: 25 May 2021

9. DOCUMENTATION OF STATISTICAL METHODS

The statistical analysis plan (version 1.0) is provided.

STATISTICAL ANALYSIS PLAN
METHODS
Protocol Number 217-MDD-301

A PHASE 3, MULTICENTER, DOUBLE-BLIND, RANDOMIZED, PLACEBO-CONTROLLED STUDY EVALUATING THE EFFICACY OF SAGE-217 IN THE TREATMENT OF ADULT SUBJECT WITH MAJOR DEPRESSIVE DISORDER (PART B)

Author of SAP: [redacted]

Version: 1.0

Version Date of SAP: 25 May 2021

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

CONFIDENTIAL

This document contains confidential information. Any use, distribution, or disclosure without the prior written consent of Sage Therapeutics, Inc. is strictly prohibited except to the extent required under applicable laws or regulations. Persons to whom the information is disclosed must be informed that the information is confidential and may not be further disclosed by them.

DocuSign Envelope ID: 93C120A4-F6C4-4D54-B80A-5E67D8DB5FBD

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:

DocuSigned by:
[redacted]
F44269FE08F64A8389AD3AB3702646A0

25-May-2021 | 16:31 EDT
Date

Approved by:

DocuSigned by:
[redacted]
Signer Name: [redacted]
Signing Time: 25-May-2021 | 17:39 EDT
18B637C692D5461B90CCB2CA3A540F5B

25-May-2021 | 17:40 EDT
Date

Sage Therapeutics, Inc.

DocuSigned by:
[redacted]
Signer Name: [redacted]
Signing Time: 26-May-2021 | 15:30 EDT
89BBDD29A7B44DDA8896FB74228A3343

26-May-2021 | 15:30 EDT
Date

Sage Therapeutics, Inc.

DocuSigned by:
[redacted]
Signer Name: [redacted]
460ADEB3CBF64E9A90A9AEF133BC820F

25-May-2021 | 16:39 EDT
Date

Sage Therapeutics, Inc.

TABLE OF CONTENTS

LIST OF TABLES

1. LIST OF ABBREVIATIONS

Abbreviation or specialist term	Explanation
AE	adverse event
AR	autoregressive
ATC	anatomical therapeutic chemical
BMI	body mass index
CGI-I	Clinical Global Impression scale for improvement
CGI-S	Clinical Global Impression scale for severity
C-SSRS	Columbia Suicide Severity Rating Scale
ECG	electrocardiogram
eCRF	electronic case report form
EOT	end of treatment
ET	early termination
FAS	Full Analysis Set
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
ICF	informed consent form
IP	Investigational product
IRT	interactive response technology
LFT	Liver Function Tests
LLOQ	Lower limit of quantification
mFAS	Modified Full Analysis Set
MADRS	Montgomery-Åsberg Depression Rating Scale
MDD	major depressive disorder
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

Abbreviation or specialist term	Explanation
[redacted]	[redacted]
PRO	patient-reported outcome
PT	preferred term
PWC-20	20-item Physician Withdrawal Checklist
SAE	serious adverse event
SAP	statistical analysis plan
SD	standard deviation
SE	standard error
SF-36	36-item Short Form survey
SI	International System of Units
SOC	System Organ Class
TEAE	treatment-emergent adverse event
UN	unstructured

2. INTRODUCTION

This statistical analysis plan (SAP) is for the final analysis of Part B data from 217-MDD-301 study, and is based on clinical study protocol, version 8.0, dated 14 December 2020.

Protocol 217-MDD-301 is a single protocol which is being conducted in 2 separate non-overlapping parts – Part A and Part B. The study protocol version 6.0 is amended after the participants enrollment was completed in Part A. Part B will enroll unique participants.

The analysis methods for data from Part A of the study are described separately, in the SAP that was approved on 12 November 2019, prior to unblinding the Part A database.

The purpose of this SAP is to describe in detail the statistical methodology and analyses to be conducted for Part B of 217-MDD-301 protocol. The SAP will be finalized and approved before database lock for Part B.

Any changes made to this SAP after the clinical database lock has occurred will be documented and discussed in the clinical study report for this study.

3. STUDY OBJECTIVES

3.1. Primary Objective

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

3.2. Secondary Objective

The secondary objective of Study 217-MDD-301 Part B is to assess patient-reported outcome (PRO) measures as they relate to health-related quality of life and depressive symptoms.

3.3. Safety Objective

The safety objective of Study 217-MDD-301 Part B is to evaluate the safety and tolerability of SAGE-217.

[redacted]

4. STUDY ENDPOINTS

4.1. Efficacy Endpoints

4.1.1. Primary Efficacy Endpoint and Estimand

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 treatment difference in mean change from baseline(see Section 8.3.3.1 for more details).

4.1.2. Secondary Efficacy Endpoints

4.1.2.1. Key Secondary Efficacy Endpoints

- Change from baseline in CGI-S at Day 15
- Change from baseline in HAM-D total score at Day 8, Day 3, and Day 42

4.1.2.2. Other Secondary Efficacy Endpoints

- HAM-D response at Day 15 and Day 42
- HAM-D remission at Day 15 and Day 42
- CGI-I response, defined as “much improved” or “very much improved”, at Day 15
- Change from baseline in MADRS total score at Day 15
- Change from baseline HAM-A total score at Day 15
- Time to first HAM-D response
- Change from baseline in PRO measures of health-related quality of life, as assessed by responses to SF-36v2, and of depressive symptoms, as assessed by the PHQ-9

4.2. 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)

[redacted]

5. STUDY DESIGN

5.1. Overall Design

Study 217-MDD-301 is a randomized, double-blind, parallel-group, placebo-controlled study in participants with major depressive disorder (MDD). This study will be conducted in 2 parts – Part A and Part B. Part B will commence after all participants in Part A have completed the Day 42 visit. Part A and Part B will enroll unique participants. For both study parts, the diagnosis of MDD must be made according to Structured Clinical Interview for Diagnostic and Diagnostic and Statistical Manual of Mental Disorder, fifth Edition for Clinical Trial (SCID-5-CT) performed by a qualified healthcare professional.

Part B

Part B will consist of a Screening Period of up to 28 days, a 14-day Treatment Period, and a 28-day double-blind follow-up period.

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. At the time of providing informed consent for the study, participants will also be asked to authorize that their unique participant identifiers be entered into a registry (www.subjectregistry.com) with the intent of identifying participants who may meet exclusion criteria due to participation in another clinical study. Participants will undergo preliminary screening procedures at the Screening Visit to determine eligibility, including completion of the HAM-D and CGI-S.

Antidepressants are permitted, provided participants are on a stable dose for at least 60 days prior to Day 1 and agree 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 will not be allowed between screening and completion of the Day 42 assessments.

Eligible participants 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 2 treatment groups (SAGE-217 50 mg or matching placebo) in a 1:1 ratio. Participants will self-administer investigational product (IP) once daily at approximately 8 PM with fat-containing food (e.g., within 1 hour of an evening meal which contains fat or with a fat-containing snack), on an outpatient basis, for 14 days. Participants will return to the study center during the treatment and follow-up periods as outlined in Section 11.1 Appendix A.

During the treatment period, participants will be able to receive IP as long as there are no dose limiting safety/tolerability concerns. Participants who cannot tolerate 50 mg will receive 40 mg for the remainder of the treatment period. At the discretion of the Investigator, participants who cannot tolerate the 40-mg dose may be discontinued from IP.

Participants who discontinue treatment early should return to the site for an EOT visit as soon as possible, preferably the day after treatment is discontinued. Follow-up visits should take place as scheduled relative to the last dose of treatment (e.g., if a participant's last dose is on Day 13, their first follow-up visit, Visit 7, should occur 4 days later). If at any time after

the EOT visit, a participant decides to terminate the study, the participant should return for an ET visit. The EOT and ET visits can be on the same day if a participant discontinues IP and decides to terminate the study on the same day during a clinic visit; in this case, all EOT visit assessments should be conducted in addition to an abbreviated physical exam.

5.2. Sample Size and Power

Using a two-sided alpha level of 0.05, a sample size of 216 evaluable participants 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 a standard deviation (SD) of 9 points. Assuming a 10% dropout rate and a 1:1 randomization ratio within each stratum (antidepressant use at baseline, yes or no), approximately 240 total randomized participants will be required to obtain 216 evaluable participants. Evaluable participants are defined as those randomized participants who receive IP and have a valid baseline and at least 1 postbaseline HAM-D assessment.

Sample size for Part B of this study was adjusted (i.e., increase) in the following amendments:

1. *In amendment #6 dated 02Mar2020, the sample size was increased to 370 to ensure enough power for subgroup analyses, and for primary efficacy endpoint analysis based on a modified Full Analysis Set (mFAS, baseline HAM-D total score >=26) that was under consideration and was included in the draft SAP version dated 14 Sep 2020.*

2. *In amendment #7 dated 14Dec2020, the sample size was further increased to approximately 575 randomized participants to provide more power for the analyses of key secondary endpoints. The increase in sample size also ensured enough power (at least 90%) for the change in primary analysis set from Full Analysis Set (FAS) to modified Full Analysis Set (mFAS) consisting all participants in the FAS with baseline HAM-D total score >=26 (mFAS is estimated to be around 60% of FAS) and to enhance comparative statistical analyses for some key subgroups. Naturally, this would also result in a higher number of subjects exposed to the SAGE-217 dose level.*

It is to note that the protocol was amended to increase the sample size to up to 575 before study enrollment close out activities were initiated 18 February 2021.

5.3. Randomization

This is a randomized, double-blind, placebo-controlled study. Participants 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; In Part B, randomization will be done within each stratum in a 1:1 ratio to receive SAGE-217 50 mg or matched placebo.

Randomization schedules will be generated by an independent statistician. The allocation to treatment group (SAGE-217 50 mg, or placebo) will be based on the randomization schedule. Randomization will be performed centrally via an interactive response technology (IRT) system.

5.4. Blinding and Unblinding

This is a randomized, double-blind, placebo-controlled study. The Sponsor, site personnel and participants will remain blinded to treatment assignment until database lock which is scheduled to occur when all participants have completed the Day 42 visit.

During the study, the blind is to be broken only when the safety of a participant is at risk and the treatment plan is dependent on the IP received. Unless a participant is at immediate risk, the Investigator should make diligent attempts to contact Sage prior to unblinding the IP administered to a participant. Requests from the Investigator about the treatment administered to study participants should be discussed with the Sage Medical Monitor. If the unblinding occurs without Sage's knowledge, the Investigator must notify Sage within 24 hours of breaking the blind.

All circumstances surrounding a premature unblinding must be clearly documented in the source records.

In all cases where the investigational product allocation for a participant is unblinded, pertinent information (including the reason for unblinding) must be documented in the participant's records and on the eCRF.

If a participant(s) or site personnel become unblinded to participant's treatment assignment before database lock, participant(s) will be excluded from the Per Protocol Set (See Section 7.5) but included in the Full Analysis Set (See Section 7.1).

6. MODIFICATIONS

6.1. Modifications to the Approved Clinical Study Protocol

Protocol 217-MDD-301, Amendment #5, 02 March 2020 introduced the Part B of 217-MDD-301 study for which this SAP is designed. The following key modifications from this initial protocol have been made:

Protocol 217-MDD-301, Amendment #6, 14 September 2020

1. *Increased sample size from 240 to approximately 370 randomized participants to ensure enough power for subgroup analyses.*
2. *Revised timepoints for the key secondary endpoint, change from baseline in HAM-D total score, to coincide with those defined in the statistical analysis plan*
3. *Added COVID-19 questions*
4. *Removed compound symmetry as a method to achieve convergence in the event of a convergence issue and specified that the sandwich estimator for the variance covariance matrix will be derived when the covariance structure is not UN*
5. *Removed score of '0 (not assessed)' as a choice for the Clinical Global Impression Scale*

Protocol 217-MDD-301, Amendment #7, 14 December 2020

1. *Increase sample size from approximately 370 to up to 575 randomized participants to ensure more power for the analyses of secondary endpoints.*
2. *Added a Modified Full Analysis Set (mFAS) to the data analysis sets and details of the estimand for the primary efficacy analysis*.

Subsequent SAP reviews by FDA led to change in protocol specified target population – it has been changed from baseline HAM-D total score >=26 to baseline HAM-D total score ≥ 24. The primary analysis will use Full Analysis Set instead of Modified FAS mentioned in the protocol.

This decision was made well in advance of database lock and unblinding of the database.

6.2. Modifications to the Approved Statistical Analysis Plan

Not applicable

6.3. Modifications to the Approved DMC Charter

Not applicable

7. ANALYSIS SETS

7.1. Full Analysis Set

The Full Analysis Set (FAS) is defined as all randomized participants administered IP with valid baseline HAM-D total score and at least 1 valid post-baseline HAM-D total score.

7.2. Modified Full Analysis Set

The Modified Full Analysis Set (mFAS) is defined as all participants in the FAS with a total HAM-D score ≥ 26 at baseline.

7.3. Safety Set

The Safety Set is defined as all participants who are administered IP.

7.4. Randomized Set

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

7.5. Per Protocol Set

The Per Protocol Set is defined as all participants in the FAS without any major protocol deviations related to efficacy. For further details, see Section 8.2.2.

In addition, Per Protocol Set will also exclude FAS participant with any of the following conditions:

1. Participants who consumed <22 capsules (i.e. <80% of assigned number of capsules)
2. Participants who consumed incorrect IP (i.e. IP other than which a participant was randomized to receive) at any time during the study
3. Participants or study personnel who were unblinded to participant's treatment assignment before database lock

[redacted]

8. STATISTICAL ANALYSIS

8.1. 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. All parameters will be converted to the International System of Units (SI) before analysis.

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 participants 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: placebo or SAGE-217. The SAGE-217 treatment group includes participants that received the 50 mg starting dose as well as those that received any 40 mg dose as a result of dose reduction. Efficacy data are analyzed by the treatment group that the participant is randomized to. Safety data are analyzed by the actual treatment received, and this is determined as follows: if a participant received any dose of SAGE-217 at any point of time, the participant is assigned to actual treatment of SAGE-217.

All participant data, including those derived, to support tables and figures will be presented in the participant data listings. In general, the participant data listings will be sorted by participant number and assessment visit and date (and time, if applicable). The treatment will be identified for each participant – either planned or actual – according to the analysis set used.

For the purpose of all safety and efficacy analyses, baseline is defined as the last non-missing measurement including unscheduled visits prior to the first dose of investigational product,

unless stated otherwise. If the time of an assessment is collected, baseline will be the latest assessment prior to first dose of IP 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 "predose".

8.1.1. Study Day Definition

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

Study day will be defined as follows:

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

8.1.2. Missing Data

All participants will be used in the analyses, as per the analysis populations, using all non-missing data available. Sensitivity analyses will be conducted to assess the impact of missing data on efficacy analysis findings. Handling of missing or incomplete dates is discussed in Section 11.3 Appendix C. Unless otherwise specified, no imputation of missing data is planned for safety analyses.

8.2. Background Characteristics

All summary tables in this section will be presented for placebo, SAGE-217, and Overall.

8.2.1. Participant Disposition

Participant disposition will be based on all participants who provided written inform consent to the study.

The summaries of participant disposition will include the number of participants who were

- screened,
- randomized,
- received IP.
- number and percentage of participants who completed the study,
- number and percentage of participants who prematurely withdrew from the study and primary reasons for not completing the study, who completed treatment,
- number and percentage of participants who discontinued treatment prematurely, and primary reasons for discontinuing treatment.

All percentages will be calculated based on the participants randomized and administered IP; treatment group assignment will be as planned treatment. If a participant is rescreened because the participant was a screen failure the first time, the status of the participant will be determined from the second screening. In the count of screened participants, this participant will be counted only once.

A completer for the study is defined as one who completed the last follow up visit (Day 42) based on the study completion CRF page with the completion question answered Yes.

A participant is marked as completing the treatment if the complete treatment question on the study treatment completion CRF page is answered Yes.

The number of participants in each analysis set will be provided. Using Randomized Set, the reason for not being included in other analysis sets will be summarized.

The summary of participant disposition will also be provided using FAS as well as mFAS.

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

8.2.2. Protocol Deviations

Protocol deviations identified during the study will be captured and categorized by the study team review as major or minor deviations in blinded fashion on an ongoing basis until database lock.

Study team will identify the major protocol deviations could affect efficacy to determine the participants in FAS to be excluded from the Per Protocol Set prior to database lock in a blinded fashion (Some major protocol deviations may not lead to participants' data to be excluded from the Per Protocol Set).

In addition, COVID-19 related protocol deviations such as remoted telephone/video visit/assessment, home healthcare visit, missed visit/assessment, out of window visit/assessment, safety reporting, IP administration, and others will be documented and provided in a separate listing if there are 10 or more participants with COVID-19 related protocol deviations. The major protocol deviations will be summarized by randomized treatment received using FAS.

All protocol deviations – major and minor - will be included in a data listing using randomized set and randomized treatment.

Any violation of inclusion/exclusion criteria will be presented with the randomized treatment in a separate data listing using Randomized Set.

8.2.3. Demographics and Baseline Characteristics

The following analyses will be provided separately for the Safety Set (using actual treatment received), the FAS (using randomized treatment), and mFAS (using randomized treatment).

Demographic data (age at informed consent date, race, sex, 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 and overall.

Baseline subgroups will be summarized for the following categories (for definition of antidepressants, see Section 8.2.5):

- Antidepressant use at baseline: Yes, No
- Age: 18-24, 25-50, 51-64 years
- Sex: Female, Male
- Race: Black or African American, White, Other
- BMI at baseline: ≤18.4, 18.5-24.9, 25-29.9, ≥30 kg/m^2
- HAM-D total score at baseline: < 26, ≥ 26; this category will not be used for mFAS.
- COVID-19 History: COVID-19 virus test positive , Suspected COVID-19, Home isolation COVID-19, COVID-19 related hospitalization, Not impacted; participant's COVID-19 history status will be collected via medical history eCRF page, and if none is present, the level is designated as 'Not impacted'.

A separate listing for participant who were randomized under incorrect stratification of antidepressant use at baseline will be provided.

The antidepressants are identified from concomitant medication records with Anatomic Therapeutic Classification (ATC) level 3 = N06A. The stratum for antidepressant use at baseline is determined from the coding of antidepressant, comparing the start/end dates of the medication versus the date of first dose of IP. Any deviation from the antidepressant stratum recorded at IRT system are part of protocol deviations and will be included in the protocol deviation listing.

Diagnostic labs are part of screening. The results of the diagnostic screening tests in Table 1 will be provided in a data listing using the Safety Set.

Child-bearing potential data will be collected at screening and will be listed in a data listing using the Safety Set.

Table 1: Diagnostic Screening Tests

Diagnostic		
Serum	**Urine**	**Breathalyzer**
Hepatitis B	Drug screen: including amphetamines, barbiturates, benzodiazepines, cannabinoids, cocaine, opiates, phencyclidine	Alcohol

Diagnostic		
Hepatitis C	Female participants 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 participants that are not surgically sterile and do not meet the protocol-defined criteria for being post-menopausal: serum hCG		
Female Participants, if menopause is suspected and not surgically sterile: FSH		

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

8.2.4. Medical/Surgical History

The following analyses will use the Safety Set.

Years since initial diagnosis of MDD, lifetime antidepressant 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 IP – date of interest. For imputation of incomplete dates in medical history, please see Section 11.3. Medical/surgical history collected at screening will be coded using the Medical Dictionary for Regulatory Activities (MedDRA), Version 23.0 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.

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

8.2.5. Prior and Concomitant Medications

The following analyses will use the Safety Set.

All medications taken 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. All medications will be coded using World Health Organization-Drug (WHO) Global B3 March 2020 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 IP. Concomitant medications are defined as those with a start date on or after the first dose of IP or those with a start date before the first dose of IP that are

ongoing or with a stop date on or after the first dose of IP. 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 11.3. Note that medication taken before the initial dosing of investigational product and continued after the initial dosing will be categorized as a prior medication and separately as a concomitant medication.

Concomitant medications will be further divided by usage period as follows:

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

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. 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 IP are permitted if the participant 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. Antidepressant medications are identified by ATC3 level code of N06A. A summary of antidepressant use at baseline and any change in these medications post-baseline (including the follow up period) will be provided. Follow-up period antidepressants are defined as those starting after the last dose of IP, but on or before the last IP dose date + 28 days.

Prohibited medications are reviewed by 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.

8.2.6. Concomitant Procedures

Concomitant procedures undergone during the study will be recorded on a separate eCRF page; this will be presented in a listing by participant. The study day for the end date of the procedure will be provided when a complete end date is available.

8.2.7. Investigational Product Exposure

The following analyses will use the Safety Set.

Total drug exposure (in mg) is defined as the total mg for SAGE-217 that was taken during the study. Total drug exposure for participants randomized to placebo is zero, unless the participant has taken SAGE-217 by mistake due to being provided wrong IP kits, in which

case the total exposure comes from SAGE-217 exposure. If the participant skips the dose on any of the days, the dose taken is 0 mg.

The kit for 50 mg dose contains two capsules – one for 30 mg, the other for 20 mg. For participants who took only 1 capsule inadvertently for a dosing day of 50 mg, it is assumed that the participant took the higher dose (i.e. 30 mg) for the day and will be calculated as such for the total drug exposure and percent of the planned exposure received. For participants who took more than 2 capsules for a dosing day of 50 mg, it assumed that the participant took all capsules as the higher dose for the day and will be calculated as such for the total drug exposure and percent of the planned exposure received.

The kit for 40 mg dose contains two capsules, 20 mg each; therefore, taking one capsule will unambiguously be assigned to 20 mg.

Total exposure duration to IP (in days) is defined as: date of last dose – date of first dose + 1. Note that this may include days when a dose is missed.

Percent of the planned exposure received is defined as the total drug exposure, divided by planned exposure, times 100. For participants randomized to SAGE-217 who complete the treatment period without dose reduction, planned exposure is 14 days of treatment planned, times 50 mg. For participants randomized to SAGE-217 who reduce dose to 40 mg, the planned exposure is the number of days (last date of 50 mg – first dose date + 1), times 50 mg, plus the number of days (date of study day 14 – first date of 40 mg + 1), times 40 mg. For participants randomized to SAGE-217 who discontinue the treatment early and reduce dose, the planned exposure is the number of days (last date of 50 mg – first dose date + 1) times 50 mg, plus the number of days (last date of dose - first date of 40 mg + 1) times 40 mg. For participants randomized to SAGE-217 who discontinued the treatment early without dose reduction, the planned exposure is the number of days (last date of 50 mg – first dose date + 1) times 50 mg. For participants 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 participants with less than 22 capsules consumed will be provided.

The number and percent of participants who had dose adjustment (reduced, interrupted, withdrawn) and dose modification (missed dose, took an extra day of dosing, took more than the planned dose for dosing day, partial dose, other) will also be summarized descriptively.

IP will be self-administered by participant once daily at approximately 8 PM with fat-containing food. Whether the participant has eaten fat-containing food within the last 1 hour of the IP administration will be included in the IP exposure data listing.

8.2.8. Investigational Product Adherence

The following analyses will use the FAS and mFAS. IP adherence (%) is defined as the number of capsules taken, divided by the number of capsules planned to be taken, times 100.

The schedule of IP is two capsules per day. The number of capsules planned to be taken is defined as follows:

1. If the participant discontinues treatment within Day 2 and Day 14 (both inclusive), the number of capsules planned to be taken is the last dose day of IP, times 2.
2. If the participant does not discontinue treatment, the number of capsules planned to be taken is 28.

IP adherence will be summarized descriptively. Number and percentage of participants with IP adherence in categories - <80%, 80-100%, >100% - will be provided.

8.3. Efficacy Analysis

8.3.1. Definition of Efficacy Variables

The efficacy variables are defined as follows:

8.3.1.1. Hamilton Rating Scale for Depression (HAM-D)

The 17-item HAM-D will be used to rate the severity of depression in participants already diagnosed as depressed. HAM-D is collected during the clinic visit on Screening, Days 1, 3, 8, 12, 15, 21, 28, 35, and 42. 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 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. Table 2 describes the subscale score calculation.

HAM-D response will be defined as having a 50% or greater reduction from baseline in HAM-D total score; only participants 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”).

Table 2: HAM-D Subscale 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 Insight	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.

8.3.1.2. 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 participant's condition post-treatment. The Investigator will rate the participant's total improvement. Response choices include: 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, 21, 28, 35, and 42. 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.

8.3.1.3. 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 participant's illness at the time of assessment, relative to the clinician's past experience with participants who have the same diagnosis. Considering total clinical experience, a participant 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, Days 1, 3, 8, 12, 15, 21, 28, 35, and 42.

8.3.1.4. 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, 28, and 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 for the HAM-A total score.

8.3.1.5. Montgomery-Åsberg Depression Rating Scale (MADRS)

The MADRS is a 10-item diagnostic questionnaire used to measure the severity of depressive episodes in participants with mood disorders. MADRS is collected at Screening, during the clinic visit on Days 1, 8, 15, 28, 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 2 individual items are missing, the MADRS total score will not be calculated and will be left as missing. If less than or equal to 2 individual item scores are missing, the missing item scores will be imputed by the mean of all other available item scores for the MADRS total score.

MADRS response will be defined as having a 50% or greater reduction from baseline in MADRS total score; only participants 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.

8.3.1.6. 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, and 42.

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 a 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 participant at each assessment, which will be used for analyses.

8.3.1.7. Patient Health Questionnaire (PHQ-9)

The PHQ-9 is a 9-item participant-rated depressive symptom severity scale. It is collected during the clinic visit on Days 1, 8, 15, 28, and 42. 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. 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.

8.3.2. Visit Windows

The scheduled visits will not be windowed and will be used at nominal visit value for treatment period visits (Day 3, Day 8, Day 12, and Day 15). For post-treatment period visits (Day 21, Day 28, Day 35, and Day 42) including unscheduled, end-of-treatment (EOT) and early termination (ET) visit will be mapped to a scheduled visit for analysis.

The windows outlined in the Table 3 for each scheduled visit have been widened compared to protocol-specified operational window, to have no gap between them; these windows are used for analysis purposes only. Once analysis visit windows are assigned, all visits, including scheduled visits, and windowed visits will be eligible for being flagged as the "analyzed record" within the analysis window; a participant'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 within treatment period, then the scheduled visit data will be used.
- If no data from the scheduled visit within treatment period is available, then the windowed visit data will be used.

- If more than one visit is in the same window, the data closest to target day will be used.
    - If there is a tie with distance from target days, the data after the target day will be used.
    - If there is a tie within a target day, the data measured at the later time will be used.

The summary by visit will use the "analyzed records" only – at most one per participant. The data not flagged as the "analyzed record" will be included in listings. A windowed visit does not fall under any analysis window will remain in the database and will be included in the listings.

Table 3 displays windows for efficacy analysis.

Table 3: Visit Windows for Efficacy Analysis

Scheduled Visit (+/1 window days) in protocol	**Target Study Day**	**Study Day Window for Visit in Analysis**
Day 1	Day 1 (pre-dose)	Day 1 (pre-dose) or last non-missing assessment before 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 (±1 day)	Day 21	Day 18 - Day 23
Day 28 (±3 day)	Day 28	Day 24 - Day 31
Day 35 (±3 days)	Day 35	Day 32 – Day 38
Day (±3 days)	Day 42	>=Day 39

8.3.3. Analysis of Efficacy Variable(s)

The FAS will be used for primary efficacy analyses. Participants will be analyzed according to randomized treatment. The mFAS will be considered as a subgroup and used for efficacy analyses for further understanding of data.

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
- HAM-D response – missing response counted as No response
- HAM-D remission
- 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
- MADRS remission
- SF-36v2 domain/component score– observed, change from baseline, percent change from baseline
- PHQ-9 score – observed (including categories), change from baseline

The HAM-D change from baseline in total score along with model-based estimates will also be presented by the following subgroups:

- Antidepressant use at baseline: yes, no
- Age group: 18-24, 25-50, 51-64 years
- Sex: Male, Female
- Race: White, Black, or African American, Other
- BMI at baseline: ≤18.4, 18.5-24.9, 25-29.9, ≥30 kg/m^2
- HAM-D total score at baseline: < 26, ≥ 26; this category will not be used for mFAS.
- COVID-19 History: COVID-19 virus test positive, Suspected COVID-19, Home isolation COVID-19, COVID-19 related hospitalization, Not impacted.

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 total score will also be presented in histogram over scheduled visits: ≤7, >7 to ≤15, >15 to ≤ 23, > 23 to <26, >=26. 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. In addition, line plots for LS means for HAM-D total score, and bar charts for HAM-D response and remission will be provided by antidepressant use at baseline.

HAM-D total scores will be also summarized for Per Protocol Set.

8.3.3.1. Mixed Effects Model for Repeated Measures

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

1) The treatment regimen for participants are: placebo or SAGE-217 for 14 days.
2) The target population is adult participants with a current diagnosis of MDD. marked by baseline HAM-D total score >=24.
3) The variable of interest is the 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 with by a sensitivity analysis using multiple imputation technique in Section 8.3.3.3.
    b. Certain medications including, but not limited to, new antidepressants and 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. Note that no rescue process is specified in the protocol, hence there is no rescue medication to be considered.
5) 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.

Data from SAGE-217 group versus placebo group will be analyzed using a mixed effects model for repeated measures (MMRM); the model will include treatment (SAGE-217 or placebo), baseline HAM-D total score, antidepressant use at baseline (Yes or No), assessment time point, 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. Model-based point estimates i.e., treatment difference in least squares [LS] mean, is the estimate of the effect 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-participant errors. If there is a convergence issue with the unstructured covariance model, Toeplitz, or Autoregressive (1) [AR (1)] covariance structure will be used, following this sequence until convergence is achieved. If the model still does not converge with AR (1) structure, no results will be reported. When the covariance structure is not UN, the sandwich estimator for the variance covariance matrix will be derived, using the EMPIRICAL option in the PROC MIXED statement in SAS. The p-value will be interpreted at 0.05 level of significance.

If the comparison of SAGE-217 versus placebo is significant at a two-sided 0.05 level, the hypotheses testing for the key secondary endpoints in FAS will proceed with multiplicity adjustment, as described in Section 8.3.3.5.

Similar to those methods described above for the primary endpoint, 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, HAM-A individual item score, MADRS total score, MADRS individual item scores, SF-36v2 Domain/Component Score, and PHQ-9 total score.

For each model, the primary comparison of interest will be between SAGE-217 and placebo at the 15-day time point. In addition, model-based point estimates (i.e., LS means), 95% confidence intervals, and p-values will be reported for the primary time point (Day 15) and all other time points.

The MMRM analysis will also be provided for change from baseline in HAM-D total score within each baseline subgroup level in Section 8.2.3 separately. If any treatment group for any level of subgroup has <=5 participants, the subgroup level will not be used in the analysis.

Line plot of model-based LS Mean and standard error (SE) over time will be presented for change from baseline in HAM-D total score, CGI-S score, HAM-A total score, and MADRS total score. Forest plot for subgroup analysis for change from baseline in HAM-D total score at Day 15 – LS mean, 95% confidence interval, and P-value – will be provided. In addition, a Forest plot for individual items in HAM-D and for subscales of HAM-D will also be provided.

SF-36v2 domain/component score– observed, change from baseline, percent change from baseline – will be summarized. Bar graphs of SF-36 domain/component score for day 15 will be provided by treatment.

PHQ-9 total score -observed, change from baseline, and total score category – will be summarized.

All efficacy analyses will be performed on FAS and also repeated on mFAS.

Summary of HAM-D model-based estimates will be provided for Per Protocol Set.

8.3.3.2. 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, or Placebo), baseline HAM-D score, antidepressant use at baseline (Yes or No), assessment time point, and time point-by-treatment as explanatory variables.

An unstructured (UN) covariance structure will be used to model the within-participant errors. If there is a convergence issue with the unstructured covariance model, then exchangeable covariance structure will be used, following this sequence until convergence is achieved. If the model still does not converge with exchangeable structure, no results will be reported.

Model-based point estimates (i.e., odds ratios), 95% confidence intervals, and p-values will be reported. These analyses will use both FAS and mFAS.

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.

Similarly, 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.

8.3.3.3. Sensitivity Analysis

A sensitivity analysis will be performed to investigate the impact of missing data if ≥ 5% of participants 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 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 FAS will be used for sensitivity analyses. Sample SAS codes for MI imputation is provided in Section 11.2, Appendix B.

Imputation distribution:

All randomized participants will be classified as non-missing category, missing category 1, or missing category 2, based on the following rules:

- Non-missing category: Participant with non-missing change from baseline in HAM-D total score at Day X
- Missing category 1: Participant discontinued due to adverse events, physician decision, protocol deviation, non-compliance with IP, or other, and is missing change from baseline in HAM-D total score at Day X
- Missing category 2: Participant discontinued due to pregnancy, lost to follow-up, participant decision, withdrawal by participant, sponsor decision, or with missing assessment due to COVID-19, and is missing change from baseline in HAM-D total score at Day X, or participant completed study but is missing change from baseline in HAM-D total score at Day X.

Imputation algorithm:

Missing values of HAM-D total score at all visits will be imputed using the fully conditional specification models. Two different imputation models will be used based on reasons of missing:

- **Missing category 1:** simulate missing values of HAM-D total score using an imputation model based on the non-missing HAM-D total scores for placebo group. This represents a conservative approach as it tends to reduce the difference between treatment and placebo group since higher values of change from baseline represents worse outcome, and placebo is supposed to provide a higher value.

- **Missing category 2:** simulate missing values of HAM-D total score using an imputation model based on the non-missing HAM-D total scores within the same treatment group.

Analysis model:

The complete multiple imputation method is described below:

- Repeat the process K (K=100) times, using the procedure described above to form K imputed complete datasets with the same variance structure.
- Fit the MMRM model including treatment, baseline antidepressant use, baseline HAM-D total score, assessment time point, and time point-by-treatment 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 $V = (K-1)\left(1 + \frac{W}{B}\right)^2$.

8.3.3.4. Analysis of Time to First HAM-D Response/Remission

Using the 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 participant will be censored at the participant's last day of HAM-D evaluation in the database if the participant did not have a response. Similar analysis will be done for first HAM-D remission.

These analyses will be also provided on mFAS.

8.3.3.5. Multiplicity Adjustment for Key Secondary Endpoints

Multiplicity adjustment to statistical testing of hypotheses of the key secondary endpoint analyses in FAS is conducted by using pre-specified fixed sequence strategy (Dmitrienko et al, 2013). Only if the primary endpoint in FAS is statistically significant at α level, the key secondary endpoints in FAS will be tested sequentially, testing each endpoint at α level of significance only if the previous endpoint in the below sequence has been significant at α level. If an endpoint is not significant at α level, formal testing of hypothesis stops and the next endpoint in the sequence will be interpreted only with nominal p-value.

The sequence of testing key secondary endpoints in FAS is as follows:

- Change from Baseline in CGI-S at Day 15
- Change from Baseline in HAM-D total score at Day 8
- Change from Baseline in HAM-D total score at Day 3
- Change from Baseline in HAM-D total score at Day 42

Other secondary endpoints not included above are not adjusted for multiplicity, and hence will be interpreted with nominal p-value.

8.3.4. Characterization of Durability of SAGE-217 Treatment Effect

After completion of 14 days of treatment, the study participant is followed for 28 days, without any further treatment with study drug. During this 28-day follow-up period, clinic visits are scheduled for Day 21, 28, 35 and 42. Day 42 change from baseline in HAM-D total score is pre-specified as a key secondary endpoint in this study.

Demonstration of SAGE-217 durable effect will be characterized via 2 complementary approaches:

1) clinically meaningful durable SAGE-217 treatment effect at Day 42 (*see Section 8.3.4.1 below*), and

2) statistically significant change from baseline at Day 42 in SAGE-217 versus placebo *(for statistical demonstration of durable SAGE-217 treatment effect, see Section 8.3.3.1 and Section 8.3.3.2 above, and Section 8.3.4.2 below*).

8.3.4.1. Durability of Clinically Meaningful Treatment Effect for SAGE-217

Durability of treatment effect is assessed over the post-treatment period in the SAGE-217 group, based on the efficacy observed at Day 15. The primary endpoint for clinically durable effect will be examined by the percent retention of the Day 15 reduction from baseline in HAM-D total score among SAGE-217 participants who had improvement in HAM-D total score at Day 15 compared to baseline; i.e. HAM-D total score at Day 15< HAM-D total score at baseline (referred to as “percent retention of Day 15 change from baseline (CFB)” going forward).

Percent retention of D15 CFB is defined as follows: Let X_b be baseline HAM-D total score, X_{15} be Day 15 HAM-D total score, X_y be Day Y (Y>15) HAM-D total score. Then percent retention (%) for Day Y is $\frac{X_y - X_b}{X_{15} - X_b} \times 100$. It will be calculated for the scheduled visits in Table 3 after Day 15, only for the participants randomized to SAGE-217. For example, for a participant with a baseline HAM-D of 27 (X_b) and a HAM-D score of 13 (X_{15}) at Day 15, percent retention at Day 42 with HAM-D of 16 (X_y) would be 79%. If Day 42 HAM-D score is 18, percent retention would be 64%.

A summary of percent retention of D15 CFB will be presented for SAGE-217 participants at Day 15 for post-Day 15 visits, based on FAS. This will serve as the primary approach for assessing clinically durable effect; mean percent retention (%) of D15 CFB at Day 42 is considered clinically meaningful durability if it is >=65%. The mean (±SE) percent retention over time will be presented in a line plot. The above analysis will also be provided for mFAS. In addition, the number and percent of participants with at least 65% retention of D15 CFB at each of post-Day 15 visit will also be provided.

Following analysis will be presented for supportive purposes to assist with further understanding of clinical durability of effect of SAGE-217. These analyses will be conducted using both FAS and mFAS.

A. A summary of percent retention of D15 CFB will be presented for SAGE-217 HAM-D responders at Day 15 for post-Day 15 visits. The mean (±SE) percent retention over time will be presented in a line plot. In addition, the number and percent of participants with at least 65% retention of D15 CFB at each of post-Day 15 visit will also be provided.

B. The number and percent of HAM-D responders at post-Day 15 visits among the SAGE-217 HAM-D responders at Day 15 will be provided. A bar chart will be provided.

C. The number and percent of HAM-D remitters (HAM-D total score <=7) at post-Day 15 visits among the SAGE-217 HAM-D remitters at Day 15 will be provided. A bar chart will be provided.

D. Relapse: A relapse is defined only for SAGE-217 HAM-D responders at Day 15. A relapse is defined as having at least 2 consecutive HAM-D total score >=20 after Day 15, including the last value. The number and percent of participants with relapse will be provided.

E. Rebound: A rebound is defined for only SAGE-217 HAM-D responders at Day 15. A rebound is any HAM-D total score (after Day 15) >= Baseline HAM-D total score. The number and percent of participants with rebound will be provided.

F. The number and percent of CGI-I responders (CGI-IR) at post-Day 15 visits among the SAGE-217 CGI-I responders at Day 15 will be provided. A bar chart will be provided.

Post Day 15 HAM-D results supporting the above analyses will be listed separately for the SAGE-217 participants based on FAS.

8.3.4.2. Durability of Treatment Effect at Day 42 via Statistical Comparison of SAGE-217 Versus Placebo

These analyses have been discussed in Section 8.3.3.1 and Section 8.3.3.2 as part of the efficacy analyses with Day 42 as one of key secondary efficacy endpoints – using FAS as primary, mFAS as subgroup:

A. MMRM analysis of change from baseline at Day 42 (key secondary efficacy endpoint), comparing SAGE-217 versus placebo (LS mean, p-value and 95% CI) will be provided.
B. GEE analysis of response/remission rates comparing SAGE-217 versus placebo at Day 42 (Odds Ratio, p-value and 95% CI) will be provided.

8.4. 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 participant has been randomized; for definition of actual treatment assignment, see Section 8.1.

The safety endpoints will be summarized by scheduled visits for treatment period visits, without any windowing. For post-treatment period visits, the choice of the visit record will be following the same rule as described in Section 8.3.2. Vital signs, C-SSRS and PWC-20 are collected as Day 18 visit for which no window exists; the evaluations will be summarized as nominal visit of Day 18.

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 IP (exclusive) and up to last dose of IP + 1 days (inclusive). Last value on study is defined as the last post-baseline value after the first dose of IP.

Any value during treatment is defined as a value on or after first IP intake but on or before last dose of randomized IP + 1 day (both inclusive). Fourteen days after last dose is defined as any value after last dose of randomized IP but on or before last dose of randomized IP + 14 days.

The safety endpoints and variables considered in the summary tables for this study are summarized in Table 4.

Table 4: Safety Endpoints and Variables in the Summary Tables

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 8.4.2-8.4.4
X = to be summarized in tables
Z = to be presented in listings only

8.4.1. Adverse Events

Adverse events (AEs) are collected starting at the time of informed consent and throughout the duration of participation in the study. A treatment-emergent adverse event (TEAE) is defined as an adverse event with onset on or after the start of IP. 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 IP dosing date/time
- TEAE: AE onset date/time on or after first IP dose date/time (If an AE start date same as IP first dose date, but no time either in AE start or treatment start, then consider this AE to be a TEAE.)
- On-Treatment TEAE: AE onset date/time on or after first IP dose date/time and on or before IP last dose date + 1 day (Note that time does not matter for the end of this period.)
- Post-Treatment TEAE: AE onset date after IP last dose date +1 day (Typically, Day 16 through Day 42 – time does not matter)

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 IP and/or last dose of IP, the adverse event will be assumed to be a TEAE and a treatment period TEAE. For imputation of missing AE dates, please refer to Appendix C, Section 11.3.

All adverse events will be coded using MedDRA version 23.0 or higher.

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

- TEAE
  - On-Treatment TEAE
  - Post-Treatment TEAE
- TEAEs by maximum severity (severe>moderate>mild)
- TEAE leading to discontinuation of IP
- TEAE leading to withdrawal from the study
- Treatment-emergent Serious Adverse Event (TESAE)
- Death

Incidence of TEAEs in following categories will be provided by SOC and PT. A participant is counted only once under each SOC and PT in case of multiple occurrences of the same AE. In these tables, the display will be sorted by descending frequency of SOC in the SAGE-217 group, then of SOC in the placebo group; if 2 or more SOCs have the same frequency, the order will be alphabetical. Within each SOC, preferred terms will be sorted by the same algorithm as in SOC.

- TEAE
- On-Treatment TEAE
- Post-Treatment TEAE
- TEAEs by maximum Severity
- TEAEs by relationship
- Serious TEAEs
- TEAEs leading to discontinuation of IP
- TEAEs leading to withdrawal from the study

Listing of AEs with onset prior to first dose of IP will be provided. All listings on TEAEs will provide 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.

For maximum severity, participants 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. A participant will be counted only once within each SOC and PT at the strongest relationship to IP in the following order: related > not related. Adverse events with onset before the first dose of IP will be provided in a separate listing. Separate data listing for deaths, SAEs, and participants who had AEs leading lose reduction will be provided.

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

- Antidepressant use at baseline: yes, no

- Age group: 18-24, 25-50, 51-64 years
- Sex: Male, Female
- Race: White, Black, or African American, Other
- BMI at baseline: ≤18.4, 18.5-24.9, 25-29.9, ≥30 kg/m^2
- Baseline HAM-D total score: < 26, ≥ 26
- COVID-19 History: COVID-19 virus test positive, Suspected COVID-19, Home isolation COVID-19, COVID-19 related hospitalization, Not impacted

8.4.2. Clinical Laboratory

The clinical laboratory tests to be performed for monitoring of safety are listed in Table 5. They are collected on screening day, days 1, 8, 15, 21, 28, and 42.

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 below limit of quantification (BLQ) and a lower limit of quantification (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.

Table 5: Clinical Laboratory Tests

Hematology	Serum Chemistry	Urinalysis	Coagulation
Red blood cell count Hemoglobin Hematocrit 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	Alanine aminotransferase Albumin Alkaline phosphatase 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	pH Specific gravity Protein Glucose Red blood cell Nitrite Leukocyte esterase Ketones Bilirubin Urobilinogen	Activated partial Thromboplastin time Prothrombin time International normalized ratio

Hematology	Serum Chemistry	Urinalysis	Coagulation
	Triglycerides Thyroid stimulating hormone (TSH) Reflex to free T3/T4 if TSH is abnormal		

Summary tables on lab parameters will include descriptive statistics for the observed values and changes from baseline by scheduled assessment timepoint 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 participant data listings. A shift table for these parameters from normal at baseline to high or low at any time during treatment, last value on treatment, and 14 days after the treatment. Qualitative urinalysis parameters will be summarized descriptively.

The number and percentage of participants with PCS values will be provided in separate displays in hematology, serum chemistry, liver function tests and urinalysis tests provided for any time during treatment, last value on treatment, and anytime within 14 days after last dose. Potentially clinically significant values will be identified for specific laboratory parameters as outlined in Table 6.

Liver function tests will be summarized for occurrence for the following parameters for these PCS threshold (for condition involving more than one parameter, 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 and included in AE displays.

Pregnancy test results will be listed but not summarized.

Table 6: Potentially Clinically Significant Values for Specific Laboratory Parameters

Laboratory Parameter	**Units**	**Criteria for PCS Values (Observed values)**	
		High	Low
Hematology			
Hemoglobin -male	g/L	>185	<115
Hemoglobin -female	g/L	>170	<100
Hematocrit-male	Fraction of 1	>0.55	<0.385
Hematocrit-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

Laboratory Parameter	**Units**	**Criteria for PCS Values (Observed values)**	
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

8.4.3. 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, 28, and 42. 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 participants with PCS and potentially clinically significant change (PCSC) values will be summarized for any time during treatment, last value on treatment, and anytime within 14 days after last dose. Potentially clinically significant values will be identified for vital sign parameters as outlined in Table 7.

Table 7: Potentially Clinically Significant Values and Change for Vital Sign Parameters

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			

Vital Sign	Units	Criteria for PCS Values (Observed values)		Criteria for PCSC values (Change from Baseline values)	
Possible orthostatic hypotension: supine – standing SBP and DBP	mmHg	SBP >= 20 and DBP >= 10			
Possible orthostatic hypotension: supine – standing SBP and DBP	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 and included in AE displays.

8.4.4. Electrocardiogram

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

The number and percentage of participants with PCS and PCSC values will be summarized for any time during treatment, last value on treatment, and anytime within 14 days after last dose. Potentially clinically significant values will be identified for ECG parameters as outlined in Table 8. 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.

Table 8: Potentially Clinically Significant Values and Change for ECG Parameters

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

8.4.5. Physical Examination

Physical examination is scheduled on screening, day 1 and 42/ET. 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.

8.4.6. 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, 18, 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 participant'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 are 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 are 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 IP, 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 IP; any Yes will make the baseline value as Yes.

The number and percentage of participants with at least one response of 'Yes' to any C-SSRS suicidal ideation or suicidal behavior item, as well as for Participant'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 all available answers for the 1st and 2nd assessment in suicidal ideation and all available 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 available 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.

8.4.7. 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 (e.g., 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. A higher total score indicates a greater degree withdrawal symptom.

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 for the PWC total score.

Instead of baseline, the first assessment on or after the last dose within 1 day of last dose of IP will be used as anchor for Day 18 and Day 21 results. The observed and change from the first assessment on or after last dose within 1 day of last dose will be summarized for PWC total score.

[redacted]

[redacted]

[redacted]

9. SUMMARY OF INTERIM AND DMC ANALYSES

Not applicable

10. REFERENCES

1. Dmitrienko A., D'Agostino, Sr. R.B., and Huque, M.F. Key multiplicity issues in clinical drug development. Statistics in Medicine. 2013 Mar; 32(7); 1079-1111

11. LIST OF APPENDICES

11.1. Appendix A: Schedule of Assessments

Visits	**Screening Period**	**Double-Blind, Placebo-Controlled Treatment Period**					**Follow-up Period**				
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**
Visit Number	**V1**	**V2**	**V3**	**V4**	**V5**	**V6**	**V7**	**V8**	**V9**	**V10**	**V11**
Study Procedure											
Informed Consent	X										
Duplicate Participant Check [b]	X										
Inclusion/Exclusion	X	X									
Serum FSH test [c]	X										
SCID-5-CT	X										
MGH ATRQ	X										
Demographics	X										
Medical/Family History[d]	X										
Participant training [e]	X	X									
Randomization		X									
Physical Examination [f]	X	X									X
Body Weight/Height	X					X (weight only)					X (weight only)

Visits	Screening Period	Double-Blind, Placebo-Controlled Treatment Period					Follow-up Period				
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
Visit Number	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11
Study Procedure											
Clinical Laboratory Assessments [g]	X	X		X		X		X	X		X
Drug & Alcohol Screen [h]	X	X	X	X	X	X	X	X	X	X	X
Pregnancy Test [i]	X	X				X			X		X
Hepatitis & HIV Screen	X										
[redacted]											
Vital Signs [l]	X	X	X	X	X	X	X		X		X
12-Lead ECG [m]	X	X				X					X
C-SSRS [n]	X	X	X	X	X	X	X	X	X	X	X
HAM-D[o,p]	X	X	X	X	X	X		X	X	X	X
MADRS		X		X		X			X		X
HAM-A[p]		X		X		X			X		X
CGI-S	X	X	X	X	X	X		X	X	X	X
CGI-I			X	X	X	X		X	X	X	X
SF-36v2	X	X		X		X			X		X
PHQ-9		X		X		X			X		X
PWC-20		X				X	X	X			

Visits	Screening Period	Double-Blind, Placebo-Controlled Treatment Period					Follow-up Period				
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
Visit Number	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11
Study Procedure											
[redacted]											
Investigational product Dispensation		X		X							
Investigational product Administration		X (Day 1 through Day 14)									
Investigational product Accountability/Return				X		X					X [r]
Adverse Events/SAEs[d, s]	X										
Prior/Concomitant Medications/Procedures[d, t]	X										

CGI-I = Clinical Global Impression – Improvement; CGI-S – Clinical Global Impression – Severity; C-SSRS = Columbia Suicide Severity Rating Scale; D = day; EOT = end of treatment; ET = early termination; ECG = electrocardiogram; FSH = follicle stimulating hormone; HAM-A = Hamilton Anxiety Rating Scale; HAM-D = Hamilton Rating Scale for Depression, 17-item; HIV = human immunodeficiency virus; 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.

[a] Participants 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. If necessary, the EOT and ET visits can be on the same day if a participant discontinues investigational product and terminates the study on the same day during a clinic visit; in this case, all EOT visit assessments should be conducted in addition to an abbreviated physical exam.

[b] Participants will be asked to authorize that their unique participant identifiers be entered into a registry (www.subjectregistrysubject.com) with the intent of identifying participants 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 participants that are not surgically sterile to confirm whether a female participant with ≥12 months of spontaneous amenorrhea meets the protocol-defined criteria for being post-menopausal.

[d] Information regarding diagnosis, isolation, and/or hospitalization due to COVID-19 will be documented as part of Medical History, AE collection, and prior/concomitant medication/procedure collection at Screening and throughout the study.

e.Participants will be trained on use of software applications and devices necessary for the conduct of the study by site personnel.

[f] 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 (e.g., head, eye, ear, nose, and throat; heart; lungs; abdomen; and extremities). An abbreviated physical exam includes a brief medical history followed by targeted physical exam.

[g] Safety laboratory tests will include hematology, serum chemistry, coagulation, and urinalysis.

[h] Urine toxicology for selected drugs of abuse (as per the lab manual) and breath test for alcohol.

[i] Serum pregnancy test at screening and urine pregnancy test thereafter for female participants who are not surgically sterile and do not meet the protocol-defined criteria for being post-menopausal.

[redacted]

[l] 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 participant has been resting for 5 minutes and then after approximately 3 minutes in the standing position. Vital signs may be repeated at the discretion of the Investigator as clinically indicated.

[m] Triplicate ECGs will be collected. When ECGs [redacted] sample collection occur on the same day, the 12-lead ECGs will be performed [redacted] sample collection.

[n] 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.

[o] The HAM-D is to be completed as early during the visit as possible.

[p] 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 scales will refer to the past 7 days (1 week) at all visits.

[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 participant'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.

11.2. Appendix B: Details of Statistical Methodology

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

- If type=un:

```
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;
lsmeans trtpn*avisitn /diff=all cl alpha=0.05;
** assuming trtpn=1 for the placebo, trtpn=2 for SAGE-217
run;
```

- If type=un does not converge, use type=TOPE or AR(1);

```
ods output lsmeans=estimates diff=diffs;
proc mixed data=&data empirical;
class trtpn avisitn usubjid antidep;
model chg=base trtpn avisitn trtpn*avisitn antidep / s;
repeated avisitn / subject=usubjid type=TOEP;
**If type=TOEP does not converge, use type=AR(1);
lsmeans trtpn*avisitn /diff=all cl alpha=0.05;
** assuming trtpn=1 for the placebo, trtpn=2 for SAGE-217
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 / within=avisitn 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=&data seed=xxxx nimpute=100 round=. . 1 1 1 1 1 1 1 1 1
out= fcs_reg1;
class trtp antidep;
```

```
FCS reg;
mnar model (base day3 day8 day 12 day15 day21 day28 day42/ modelobs= (trtp='A'));
var strata base day3 day8 day12 day15 day21 day28 day35 day42;
run;
** Missing category 2
proc mi data=&data seed=xxxx nimpute=100 round=. . 1 1 1 1 1 1 1 1 1 out=fcs_reg2;
class trtp strata;
fcs nbiter=20 reg (base day3 day8 day12 day15 day21 day28 day35 day42/details);
var trtp strata base day3 day8 day12 day15 day21 day28 day35 day42;
run;
```

11.3. 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.

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 min (first dose date + 1 day, last dose date).
    - If the year of AE onset < year of initiation of the treatment, then the month and day will be set to December 31st.
    - If the year of AE onset > the year of initiation of treatment, then the month and day will be set to January 1st.
- 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 of the particular year.
        - 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 of the particular year.
  - 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.

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

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 end date (imputed date if applicable) 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 min (December 31, date of death).
  - If the year and day are present and the month is missing, then the month will be set to min (December, month of death). If the year and month are present and the day is missing, then the day will be set to min (last day of the month, month of death).
  - 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 min (December 31, date of death).

11.4. Appendix E: List of Displays

Tables

Table Number	Title	Analysis Set
Table 14.1.1.1	Summary of Participant Disposition	All Participants
Table 14.1.1.2	Summary of Analysis Sets	All Participants
Table 14.1.1.3	Summary of Participant Disposition	Full Analysis Set
Table 14.1.1.4	Summary of Participant Disposition	Modified Full Analysis Set
Table 14.1.2.1	Summary of Major Protocol Deviations	Full Analysis 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 14.1.3.1.2.1	Summary of Demographics and Baseline Characteristics	Modified Full Analysis Set
Table 14.1.3.1.3	Summary of Baseline Subgroups	Safety Set
Table 14.1.3.1.4	Summary of Baseline Subgroups	Full Analysis Set
Table 14.1.3.1.5	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 Participant 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 Number	Title	Analysis 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 Investigational Product Exposure	Safety Set
Table 14.1.4.11	Summary of Investigational Product Adherence	Full Analysis Set
Table 14.1.4.12	Summary of Investigational Product 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 Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit	Full Analysis Set
Table 14.2.1.1.3	Model-based Sensitivity Analysis on Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score 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 Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score 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 Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit	Modified Full Analysis Set

Table Number	Title	Analysis Set
Table 14.2.1.1.8	Model-based Sensitivity Analysis on Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score at Day 15	Modified Full Analysis Set
Table 14.2.1.2.1	Summary of Hamilton Rating Scale for Depression (HAM-D) Subscale Scores by Study Visit	Full Analysis Set
Table 14.2.1.2.2	Model-based Results on Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Subscale Score by Study Visit	Full Analysis Set
Table 14.2.1.2.3	Summary of Hamilton Rating Scale for Depression (HAM-D) Subscale Scores by Study Visit	Modified Full Analysis Set
Table 14.2.1.2.4	Model-based Results on Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Subscale Score by Study Visit	Modified Full Analysis Set
Table 14.2.1.3.1	Summary of Hamilton Rating Scale for Depression (HAM-D) Individual Item Score by Study Visit	Full Analysis Set
Table 14.2.1.3.2	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Individual Item Score, Change from Baseline by Study Visit	Full Analysis Set
Table 14.2.1.3.3	Summary of Hamilton Rating Scale for Depression (HAM-D) Individual Item Score by Study Visit	Modified Full Analysis Set
Table 14.2.1.3.4	Model-based Results on Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Individual Item Score by Study Visit	Modified Full Analysis Set
Table 14.2.1.4.1	Summary of Hamilton Rating Scale for Depression (HAM-D) Response by Study Visit	Full Analysis Set
Table 14.2.1.4.2	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Response by Study Visit	Full Analysis Set
Table 14.2.1.4.3	Sensitivity Analysis: Summary of Hamilton Rating Scale for Depression (HAM-D) Response by Study Visit	Full Analysis Set
Table 14.2.1.4.4	Summary of Hamilton Rating Scale for Depression (HAM-D) Total Score – Percent Improvement – by Study Visit	Full Analysis Set

Table Number	Title	Analysis Set
Table 14.2.1.4.5	Summary of Time to First Hamilton Rating Scale for Depression (HAM-D) Response – Kaplan-Meier Analysis	Full Analysis Set
Table 14.2.1.4.6	Summary of Hamilton Rating Scale for Depression (HAM-D) Response by Study Visit	Modified Full Analysis Set
Table 14.2.1.4.7	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Response by Study Visit	Modified Full Analysis Set
Table 14.2.1.4.8	Sensitivity Analysis: Summary of Hamilton Rating Scale for Depression (HAM-D) Response by Study Visit	Modified Full Analysis Set
Table 14.2.1.4.9	Summary of Hamilton Rating Scale for Depression (HAM-D) Total Score – Percent Improvement – by Study Visit	Modified Full Analysis Set
Table 14.2.1.4.10	Summary of Time to First Hamilton Rating Scale for Depression (HAM-D) Response – Kaplan-Meier Analysis	Modified Full Analysis Set
Table 14.2.1.5.1	Summary of Hamilton Rating Scale for Depression (HAM-D) Remission by Study Visit	Full Analysis Set
Table 14.2.1.5.2	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Remission by Study Visit	Full Analysis Set
Table 14.2.1.5.3	Sensitivity Analysis: Summary of Hamilton Rating Scale for Depression (HAM-D) Remission by Study Visit	Full Analysis Set
Table 14.2.1.5.4	Summary of Hamilton Rating Scale for Depression (HAM-D) Total Score in Categories, by Study Visit	Full Analysis Set
Table 14.2.1.5.5	Summary of Time to First Hamilton Rating Scale for Depression (HAM-D) Remission – Kaplan-Meier Analysis	Full Analysis Set
Table 14.2.1.5.6	Summary of Hamilton Rating Scale for Depression (HAM-D) Remission by Study Visit	Modified Full Analysis Set
Table 14.2.1.5.7	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Remission by Study Visit	Modified Full Analysis Set
Table 14.2.1.5.9	Summary of Hamilton Rating Scale for Depression (HAM-D) Total Score in Categories, by Study Visit	Modified Full Analysis Set

Table Number	Title	Analysis Set
Table 14.21.5.10	Summary of Time to First Hamilton Rating Scale for Depression (HAM-D) Remission – Kaplan-Meier Analysis	Modified Full Analysis Set
Table 14.2.1.6.1	Summary of HAM-D Total Score by Study Visit and Antidepressant Use at Baseline	Full Analysis Set
Table 14.2.1.6.2	Model-based Results on Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit and Antidepressant Use at Baseline	Full Analysis Set
Table 14.2.1.6.3	Summary of HAM-D Total Score by Study Visit and Age Group	Full Analysis Set
Table 14.2.1.6.4	Model-based Results on Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit and Age Group	Full Analysis Set
Table 14.2.1.6.5	Summary of HAM-D Total Score by Study Visit and Sex	Full Analysis Set
Table 14.2.1.6.6	Model-based Results on Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit and Sex	Full Analysis Set
Table 14.2.1.6.7	Summary of HAM-D Total Score by Study Visit and Race Group	Full Analysis Set
Table 14.2.1.6.8	Model-based Results on Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit and Race Group	Full Analysis Set
Table 14.2.1.6.9	Summary of HAM-D Total Score by Study Visit and BMI at Baseline	Full Analysis Set
Table 14.2.1.6.10	Model-based Results on Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit and BMI at Baseline	Full Analysis Set
Table 14.2.1.6.11	Summary of HAM-D Total Score by Study Visit and Baseline HAM-D Total Score Category	Full Analysis Set
Table 14.2.1.6.12	Model-based Results on Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit and Baseline HAM-D Total Score Category	Full Analysis Set
Table 14.2.1.6.13	Summary of HAM-D Total Score by Study Visit and COVID-19 History	Full Analysis Set
Table 14.2.1.6.14	Model-based Results on Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit and COVID-19 History	Full Analysis Set
Table 14.2.1.6.15	Summary of HAM-D Total Score by Study Visit and Antidepressant Use at Baseline	Modified Full Analysis Set

Table Number	Title	Analysis Set
Table 14.2.1.6.16	Model-based Results on Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit and Antidepressant Use at Baseline	Modified Full Analysis Set
Table 14.2.1.6.17	Summary of HAM-D Total Score by Study Visit and Age Group	Modified Full Analysis Set
Table 14.2.1.6.18	Model-based Results on Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit and Age Group	Modified Full Analysis Set
Table 14.2.1.6.19	Summary of HAM-D Total Score by Study Visit and Sex	Modified Full Analysis Set
Table 14.2.1.6.20	Model-based Results on Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit and Sex	Modified Full Analysis Set
Table 14.2.1.6.21	Summary of HAM-D Total Score by Study Visit and Race Group	Modified Full Analysis Set
Table 14.2.1.6.22	Model-based Results on Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit and Race Group	Modified Full Analysis Set
Table 14.2.1.6.23	Summary of HAM-D Total Score by Study Visit and BMI at Baseline	Modified Full Analysis Set
Table 14.2.1.6.24	Model-based Results on Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit and BMI at Baseline	Modified Full Analysis Set
Table 14.2.1.6.27	Summary of HAM-D Total Score by Study Visit and COVID-19 History	Modified Full Analysis Set
Table 14.2.1.6.28	Model-based Results on Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score by Study Visit and COVID-19 History	Modified Full Analysis Set
Table 14.2.1.6.29	Summary of HAM-D Response by Study Visit and Antidepressant Use at Baseline	Full Analysis Set
Table 14.2.1.6.30	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Response by Study Visit and Antidepressant Use at Baseline	Full Analysis Set
Table 14.2.1.6.31	Summary of HAM-D Response by Study Visit and Antidepressant Use at Baseline	Modified Full Analysis Set
Table 14.2.1.6.32	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Response by Study Visit and Antidepressant Use at Baseline	Modified Full Analysis Set
Table 14.2.1.6.33	Summary of HAM-D Remission by Study Visit and Antidepressant Use at Baseline	Full Analysis Set

Table Number	Title	Analysis Set
Table 14.2.1.6.34	Model-based Results Hamilton Rating Scale for Depression (HAM-D) Remission by Study Visit and Antidepressant Use at Baseline	Full Analysis Set
Table 14.2.1.6.35	Summary of HAM-D Remission by Study Visit and Antidepressant Use at Baseline	Modified Full Analysis Set
Table 14.2.1.6.36	Model-based Results on Hamilton Rating Scale for Depression (HAM-D) Remission by Study Visit and Antidepressant Use at Baseline	Modified Full Analysis Set
Table 14.2.2.1.1	Summary of Clinical Global Impression – Improvement (CGI-I) by Study Visit	Full Analysis Set
Table 14.2.2.2.1	Summary of Clinical Global Impression – Improvement (CGI-I) Response by Study Visit	Full Analysis Set
Table 14.2.2.2.2	Model-based Results on Clinical Global Impression – Improvement (CGI-I) Response by Study Visit	Full Analysis Set
Table 14.2.2.3.1	Summary of Clinical Global Impression – Improvement (CGI-I) by Study Visit	Modified Full Analysis Set
Table 14.2.2.3.2	Summary of Clinical Global Impression – Improvement (CGI-I) Response by Study Visit	Modified Full Analysis Set
Table 14.2.2.3.3	Model-based Results on Clinical Global Impression – Improvement (CGI-I) Response by Study Visit	Modified Full Analysis Set
Table 14.2.3.1.1	Summary of Clinical Global Impression – Severity (CGI-S) by Study Visit	Full Analysis Set
Table 14.2.3.1.2	Model-based Results on Change from Baseline in Clinical Global Impression – Severity (CGI-S) by Study Visit	Full Analysis Set
Table 14.2.3.1.3	Summary of Clinical Global Impression – Severity (CGI-S) by Study Visit	Modified Full Analysis Set
Table 14.2.3.1.4	Model-based Results on Change from Baseline in Clinical Global Impression – Severity (CGI-S) by Study Visit	Modified Full Analysis Set
Table 14.2.4.1.1	Summary of Hamilton Rating Scale for Anxiety (HAM-A) Total Score by Visit	Full Analysis Set
Table 14.2.4.1.2	Model-based Results on Change from Baseline in Hamilton Rating Scale for Depression (HAM-A) Total Score by Study Visit	Full Analysis Set

Table Number	Title	Analysis Set
Table 14.2.4.2.1	Summary of Hamilton Rating Scale for Anxiety (HAM-A) Individual Item Score by Study Visit	Full Analysis Set
Table 14.2.4.2.2	Model-based Results on Change from Baseline in Hamilton Rating Scale for Depression (HAM-A) Individual Item Score by Study Visit	Full Analysis Set
Table 14.2.4.3.1	Summary of Hamilton Rating Scale for Anxiety (HAM-A) Total Score by Visit	Modified Full Analysis Set
Table 14.2.4.3.2	Model-based Results on Change from Baseline in Hamilton Rating Scale for Depression (HAM-A) Total Score by Study Visit	Modified Full Analysis Set
Table 14.2.4.4.1	Summary of Hamilton Rating Scale for Anxiety (HAM-A) Individual Item Score by Study Visit	Modified Full Analysis Set
Table 14.2.4.4.2	Model-based Results on Change from Baseline in Hamilton Rating Scale for Depression (HAM-A) Individual Item Score by Study Visit	Modified Full Analysis Set
Table 14.2.5.1.1	Summary of Montgomery-Asberg Depression Rating Scale (MADRS) Total Score by Study Visit	Full Analysis Set
Table 14.2.5.1.2	Model-based Results on Change from Baseline in Montgomery-Asberg Depression Rating Scale (MADRS) Total Score by Study Visit	Full Analysis Set
Table 14.2.5.2.1	Summary of Montgomery-Asberg Depression Rating Scale (MADRS) Individual Item Score by Study Visit	Full Analysis Set
Table 14.2.5.2.2	Model-based Results on Change from Baseline in Montgomery-Asberg Depression Rating Scale (MADRS) Individual Item Score	Full Analysis Set
Table 14.2.5.3.1	Summary of Montgomery-Asberg Depression Rating Scale (MADRS) Response by Study Visit	Full Analysis Set
Table 14.2.5.3.2	Model-based Results on Montgomery-Asberg Depression Rating Scale (MADRS) Response by Study Visit	Full Analysis Set
Table 14.2.5.4.1	Summary of Montgomery-Asberg Depression Rating Scale (MADRS) Remission by Study Visit	Full Analysis Set

Table Number	Title	Analysis Set
Table 14.2.5.4.2	Model-based Results on Montgomery-Asberg Depression Rating Scale (MADRS) Remission by Study Visit	Full Analysis Set
Table 14.2.5.5.1	Summary of Montgomery-Asberg Depression Rating Scale (MADRS) Total Score by Study Visit	Modified Full Analysis Set
Table 14.2.5.5.2	Model-based Results on Change from Baseline in Montgomery-Asberg Depression Rating Scale (MADRS) Total Score by Study Visit	Modified Full Analysis Set
Table 14.2.5.6.1	Summary of Montgomery-Asberg Depression Rating Scale (MADRS) Individual Item Score by Study Visit	Modified Full Analysis Set
Table 14.2.5.6.2	Model-based Results on Change from Baseline in Montgomery-Asberg Depression Rating Scale (MADRS) Individual Item Score	Modified Full Analysis Set
Table 14.2.5.7.1	Summary of Montgomery-Asberg Depression Rating Scale (MADRS) Response by Study Visit	Modified Full Analysis Set
Table 14.2.5.7.2	Model-based Results on Montgomery-Asberg Depression Rating Scale (MADRS) Response by Study Visit	Modified Full Analysis Set
Table 14.2.5.8.1	Summary of Montgomery-Asberg Depression Rating Scale (MADRS) Remission by Study Visit	Modified Full Analysis Set
Table 14.2.5.8.2	Model-based Results on Montgomery-Asberg Depression Rating Scale (MADRS) Remission by Study Visit	Modified Full Analysis Set
Table 14.2.6.1.1	Summary of Short Form-36 Version 2 (SF-36v2) Domain/Component Score by Study Visit	Full Analysis Set
Table 14.2.6.1.2	Model-based Results on Change from Baseline in Short Form-36 Version 2 (SF-36v2) Individual Domain/Component by Study Visit	Full Analysis Set
Table 14.2.6.2.1	Summary of Short Form-36 Version 2 (SF-36v2) Domain/Component Score by Study Visit	Modified Full Analysis Set
Table 14.2.6.2.2	Model-based Results on Change from Baseline in Short Form-36 Version 2 (SF-36v2) Individual Domain/Component by Study Visit	Modified Full Analysis Set

Table Number	Title	Analysis Set
Table 14.2.7.1.1	Summary of Patient Health Questionnaire (PHQ-9) Total Score by Study Visit	Full Analysis Set
Table 14.2.7.1.2	Summary of Patient Health Questionnaire (PHQ-9) Total Score Category by Study Visit	Full Analysis Set
Table 14.2.7.1.3	Model-based Results on Change from Baseline in Patient Health Questionnaire (PHQ-9) Total Score by Study Visit	Full Analysis Set
Table 14.2.7.2.1	Summary of Patient Health Questionnaire (PHQ-9) Total Score by Study Visit	Modified Full Analysis Set
Table 14.2.7.2.2	Summary of Patient Health Questionnaire (PHQ-9) Total Score Category by Study Visit	Modified Full Analysis Set
Table 14.2.7.2.3	Model-based Results on Change from Baseline in Patient Health Questionnaire (PHQ-9) Total Score by Study Visit	Modified Full Analysis Set
Table 14.2.8.1	Summary of Percent Retention (%) of Day 15 Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) at Post-Day15 Visits	Full Analysis Set-Participants who had improvement at Day 15
Table 14.2.8.2	Summary of Percent Retention (%) of Day 15 Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) at Post-Day 15 Visits	Modified Full Analysis Set-Participants who had improvement at Day 15
Table 14.2.8.3	Summary of Percent Retention (%) of Day 15 Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) at Post-D15 Visits, for HAM-D Responders at Day 15	Full Analysis Set -Day 15 HAM-D Responders Only
Table 14.2.8.4	Summary of Percent Retention (%) of Day 15 Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) at Post-D15 Visits, for HAM-D Responders at Day 15	Modified Full Analysis Set -Day 15 HAM-D Responders Only
Table 14.2.8.5	Summary of Hamilton Rating Scale for Depression (HAM-D) Response at Post-Day 15 Visits, for HAM-D Responders at Day 15	Full Analysis Set -Day 15 HAM-D Responders Only
Table 14.2.8.6	Summary of Hamilton Rating Scale for Depression (HAM-D) Response at Post-Day 15 Visits, for HAM-D Responders at Day 15	Modified Full Analysis Set -Day 15 HAM-D Responders Only
Table 14.2.8.7	Summary of Hamilton Rating Scale for Depression (HAM-D) Remission at Post-Day 15 Visits, for HAM-D Remitters at Day 15	Full Analysis Set -Day 15 HAM-D Responders Only

Table Number	Title	Analysis Set
Table 14.2.8.8	Summary of Hamilton Rating Scale for Depression (HAM-D) Remission at Post-Day 15 Visits, for HAM-D Remitters at Day 15	Modified Full Analysis Set-Day 15 HAM-D Responders Only
Table 14.2.8.9	Summary of Hamilton Rating Scale for Depression (HAM-D) Relapse and Rebound Post-Day 15, for HAM-D Responders at Day 15	Full Analysis Set-Day 15 HAM-D Responders Only
Table 14.2.8.10	Summary of Hamilton Rating Scale for Depression (HAM-D) Relapse and Rebound Post-Day 15, for HAM-D Responders at Day 15	Modified Full Analysis Set-Day 15 HAM-D Responders Only
Table 14.2.8.11	Summary of Clinical Global Impression – Improvement (CGI-I) Response at Post-Day 15 Visits, for CGI-I Responders at Day 15	Full Analysis Set-Day 15 HAM-D Responders Only
Table 14.2.8.12	Summary of Clinical Global Impression – Improvement (CGI-I) Response at Post-Day 15 Visits, for CGI-I Responders at Day 15	Modified Full Analysis Set-Day 15 HAM-D Responders Only
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 14.3.1.3.1	Summary of On-Treatment Emergent Adverse Events by System Organ Class and Preferred Term	Safety Set
Table 14.3.1.3.2	Summary of Post- Treatment-Emergent Adverse Events by System Organ Class and Preferred Term	Safety Set
Table 14.3.1.4.1	Summary of Treatment-Emergent Adverse Events by System Organ Class, Preferred Term, and Use of Antidepressant at Baseline	Safety Set
Table 14.3.1.4.2	Summary of Treatment-Emergent Adverse Events by System Organ Class, Preferred Term, and Age Group	Safety Set
Table 14.3.1.4.3	Summary of Treatment-Emergent Adverse Events by System Organ Class, Preferred Term, and Sex	Safety Set
Table 14.3.1.4.4	Summary of Treatment-Emergent Adverse Events by System Organ Class, Preferred Term, and Race Group	Safety Set

Table Number	Title	Analysis Set
Table 14.3.1.4.5	Summary of Treatment-Emergent Adverse Events by System Organ Class, Preferred Term, and BMI at Baseline	Safety Set
Table 14.3.1.4.6	Summary of Treatment-Emergent Adverse Events by System Organ Class, Preferred Term, and Baseline HAM-D Total Score	Safety Set
Table 14.3.1.4.7	Summary of Treatment-Emergent Adverse Events by System Organ Class, Preferred Term, and COVID-19 History	Safety Set
Table 14.3.1.5	Summary of Treatment-Emergent Adverse Events by System Organ Class, Preferred Term, and Maximum Severity	Safety Set
Table 14.3.1.6	Summary of Treatment-Emergent Adverse Events by System Organ Class, Preferred Term, and Relationship to Investigational Product	Safety Set
Table 14.3.1.7	Summary of Treatment Period Treatment-Emergent Adverse Events Leading to Discontinuation of Investigational Product by System Organ Class and Preferred Term	Safety Set
Table 14.3.1.8	Summary of Treatment-Emergent Adverse Events Leading to Withdrawal from the Study by System Organ Class and Preferred Term	Safety Set
Table 14.3.1.9	Summary of Treatment-Emergent Serious Adverse Events by System Organ Class and Preferred Term	Safety Set
Table 14.3.1.10	Summary of Most Common (>2%) Treatment-Emergent Adverse Events by Preferred Term	Safety Set
Table 14.3.1.11	Summary of Treatment-Emergent Adverse Events Leading to Dose Reduction by System Organ Class, Preferred Term	Safety Set
Table 14.3.4.1.1	Summary of Serum Chemistry Results by Study Visit	Safety Set
Table 14.3.4.1.2	Summary of Shift in Serum Chemistry Results	Safety Set
Table 14.3.4.1.3	Summary of Potentially Clinically Significant Liver Function Tests	Safety Set
Table 14.3.4.1.4	Summary of Potentially Clinically Significant Serum Chemistry (excluding LFT)	Safety Set
Table 14.3.4. 2.1	Summary of Hematology Results by Study Visit	Safety Set

Table Number	Title	Analysis Set
Table 14.3.4.2.2	Summary of Shift in Hematology Results	Safety Set
Table 14.3.4.2.3	Summary of Potentially Clinically Significant Hematology Results	Safety Set
Table 14.3.4. 3.1	Summary of Urinalysis Results by Study Visit – Quantitative Results	Safety Set
Table 14.3.4. 3.2	Summary of Urinalysis by Study Visit – Qualitative Results	Safety Set
Table 14.3.4.4.1	Summary of Coagulation Results by Study Visit	Safety Set
Table 14.3.4.5.1	Summary of Vital Signs by Study Visit	Safety Set
Table 14.3.4.5.2	Summary of Potentially Clinically Significant Vital Signs Any Time Post-Baseline	Safety Set
Table 14.3.4.5.3	Summary of Orthostatic Changes (Supine minus Standing in Vital Signs by Study Visit	Safety Set
Table 14.3.4.6.1	Summary of ECG data by Study Visit	Safety Set
Table 14.3.4.6.2	Summary of Abnormal ECG by Study Visit	Safety Set
Table 14.3.4.6.3	Summary of Potentially Clinically Significant QTcF in ECG Data Any Time Post-Baseline	Safety Set
Table 14.3.4.7.1	Summary of Columbia Suicide Rating Scale (C-SSRS) Suicidal Ideation and Suicidal Behavior by Study Visit and Treatment Group	Safety Set
Table 14.3.4.7.2	Summary of Columbia-Suicide Severity Rating Scale (C-SSRS) on Suicidal Ideation and Suicidal Behavior Data	Safety Set
Table 14.3.4.7.3.1	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. 7.3.2	Summary of Shift from Baseline in C-SSRS Maximum Severity Score in Suicidal Ideation at Any Time Post-Baseline	Safety Set
Table 14.3.4.8	Summary of Physical Withdrawal Checklist (PWC)-20 Total Scores by Study Visit	Safety Set

Figures

Figure Number	Title	Analysis Set
Figure 14.2.1.1	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 14.2.1.1.1	Line Plot of LS Mean (±SE) Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score Over Time by Treatment Group and Antidepressant Use at Baseline	Full Analysis Set
Figure 14.2.1.2	Forest Plot of Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score at Day 15 by Treatment Group and Subgroups	Full Analysis Set
Figure 14.2.1.3	Bar Chart of Hamilton Rating Scale for Depression (HAM-D) Response over Time	Full Analysis Set
Figure 14.2.1.3.1	Bar Chart of Hamilton Rating Scale for Depression (HAM-D) Response over Time, by Antidepressant Use at Baseline	Full Analysis Set
Figure 14.2.1.4	Histogram of Hamilton Rating Scale for Depression (HAM-D) Percentage Improvement over Time	Full Analysis Set
Figure 14.2.1.5	Bar Chart of Hamilton Rating Scale for Depression (HAM-D) Remission over Time	Full Analysis Set
Figure 14.2.1.5.1	Bar Chart of Hamilton Rating Scale for Depression (HAM-D) Remission over Time, by Antidepressant Use at Baseline	Full Analysis Set
Figure 14.2.1.6	Histogram of Hamilton Rating Scale for Depression (HAM-D) Total Score in Categories over Time	Per Protocol Set
Figure 14.2.1.7	Forest Plot of Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Subscales on Day 15 by Treatment Group	Full Analysis Set
Figure 14.2.1.8	Forest Plot of Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) for Individual Items on Day 15 by Treatment Group	Full Analysis Set
Figure 14.2.2	Bar Chart of CGI-I Response over Time by Treatment Group	Full Analysis Set
Figure 14.2.3	Line Plot of LS Mean (±SE) Change from Baseline in CGI-S over Time by Treatment Group	Full Analysis Set

Figure Number	Title	Analysis Set
Figure 14.2.4	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 14.2.5	Line Plot of LS Mean (±SE) Change from Baseline in Montgomery-Åsberg Depression Rating Scale (MADRS) Total Score over Time by Treatment Group	Full Analysis Set
Figure 14.2.6	Bar Chart of LS mean (± SE) Change from Baseline in Short Form-36 Version 2 at Day 15 by Domain/Component and Treatment Group	Full Analysis Set
Figure 14.2.7	Distribution of Participants with Suicide-Related Events Based on the C-SSRS Scores over Time by Treatment Group	Safety Set
Figure 14.2.8.1	Line Plot of LS Mean (±SE) Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score Over Time by Treatment Group	Modified Full Analysis Set
Figure 14.2.8.1.1	Line Plot of LS Mean (±SE) Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score Over Time by Treatment Group and Antidepressant Use at Baseline	Modified Full Analysis Set
Figure 14.2.8.2	Forest Plot of Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Total Score at Day 15 by Treatment Group and Subgroups	Modified Full Analysis Set
Figure 14.2.8.3	Bar Chart of Hamilton Rating Scale for Depression (HAM-D) Response over Time	Modified Full Analysis Set
Figure 14.2.8.3.1	Bar Chart of Hamilton Rating Scale for Depression (HAM-D) Response over Time, by Antidepressant Use at Baseline	Modified Full Analysis Set
Figure 14.2.8.4	Histogram of Hamilton Rating Scale for Depression (HAM-D) Percentage Improvement over Time	Modified Full Analysis Set
Figure 14.2.8.5	Bar Chart of Hamilton Rating Scale for Depression (HAM-D) Remission over Time	Modified Full Analysis Set
Figure 14.2.8.5.1	Bar Chart of Hamilton Rating Scale for Depression (HAM-D) Remission over Time, by Antidepressant Use at Baseline	Modified Full Analysis Set
Figure 14.2.8.6	Histogram of Hamilton Rating Scale for Depression (HAM-D) Total Score in Categories over Time	Modified Full Analysis Set
Figure 14.2.8.7	Forest Plot of Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) Subscales on Day 15 by Treatment Group	Modified Full Analysis Set
Figure 14.2.8.8	Forest Plot of Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) for Individual Items on Day 15 by Treatment Group	Modified Full Analysis Set

Figure Number	Title	Analysis Set
Figure 14.2.9	Bar Chart of CGI-I Response over Time by Treatment Group	Modified Full Analysis Set
Figure 14.2.10	Line Plot of LS Mean (±SE) Change from Baseline in CGI-S over Time by Treatment Group	Modified Full Analysis Set
Figure 14.2.11	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 14.2.12	Line Plot of LS Mean (±SE) Change from Baseline in Montgomery-Åsberg Depression Rating Scale (MADRS) Total Score over Time by Treatment Group	Modified Full Analysis Set
Figure 14.2.13	Bar Chart of LS mean (± SE) Change from Baseline in Short Form-36 Version 2 at Day 15 by Domain/Component and Treatment Group	Modified Full Analysis Set
Figure 14.2.14	Line Plot of Percent Retention (%) of Day 15 Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) over Time	Full Analysis Set-Participants who had improvement at Day 15
Figure 14.2.15	Line Plot of Percent Retention (%) of Day 15 Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) over Time	Modified Full Analysis Set-Participants who had improvement at Day 15
Figure 14.2.16	Line Plot of Percent Retention (%) of Day 15 Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) over Time, for HAM-D Responders at Day 15	Full Analysis Set -Day 15 HAM-D Responders Only
Figure 14.2.17	Line Plot of Percent Retention (%) of Day 15 Change from Baseline in Hamilton Rating Scale for Depression (HAM-D) over Time, for HAM-D Responders at Day 15	Modified Full Analysis Set--Day 15 HAM-D Responders Only
Figure 14.2.18	Bar Chart of Hamilton Rating Scale for Depression (HAM-D) Response at Post-Day 15 Visits, for HAM-D Responders at Day 15, for HAM-D Responders at Day 15	Full Analysis Set -Day 15 HAM-D Responders Only
Figure 14.2.19	Bar Chart of Hamilton Rating Scale for Depression (HAM-D) Response at Post-Day 15 Visits, for HAM-D Responders at Day 15, for HAM-D Responders at Day 15	Modified Full Analysis Set--Day 15 HAM-D Responders Only
Figure 14.2.20	Bar Chat of Hamilton Rating Scale for Depression (HAM-D) Remission at Post-Day 15 Visits, for HAM-D Remitters at Day 15, for HAM-D Remitters at Day 15	Full Analysis Set -Day 15 HAM-D Remitters Only

Figure Number	Title	Analysis Set
Figure 14.2.21	Bar Chat of Hamilton Rating Scale for Depression (HAM-D) Remission at Post-Day 15 Visits, for HAM-D Remitters at Day 15, for HAM-D Remitters at Day 15	Modified Full Analysis Set-Day 15 HAM-D Remitters Only
Figure 14.2.22	Bar Chart of Clinical Global Impression – Improvement (CGI-I) Response at Post-Day 15 Visits, for CGI-I Responders at Day 15, for CGI-I Responders at Day 15	Full Analysis Set-Day 15 CGI-I Responders Only
Figure 14.2.23	Bar Chart of Clinical Global Impression – Improvement (CGI-I) Response at Post-Day 15 Visits, for CGI-I Responders at Day 15, for CGI-I Responders at Day 15	Modified Full Analysis Set-Day 15 CGI-I Responders Only
Figure 14.2.24	Box Plot of Physician Withdrawal Checklist (PWC)-20 Total Score from First Assessment after Last Dose within 1 Day Past Last Dose by Study Visit	Safety Set

Listings

Listing Number	**Title**	**Analysis Set**
Listing 16.1.7	Participant Randomization	Randomized Set
Listing 16.2.1.1	Participant Disposition	Randomized Set
Listing 16.2.1.2	Premature Discontinuation from Investigational Product and/or Withdrawal from the Study	Safety Set
Listing 16.2.2.1	Protocol Deviations	Safety Set
Listing 16.2.2.1.1	Assessments Affected by COVID-19	Safety Set
Listing 16.2.2.2	Inappropriate Investigational Product 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 16.2.4.2.3	Child-Bearing Potential at Screening	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	Participant 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	Concomitant Procedures	Safety Set

Listing Number	Title	Analysis Set
Listing 16.2.5.1	Investigational Product Administration	Safety Set
Listing 16.2.5.2	Investigational Product Exposure	Safety Set
Listing 16.2.5.3	Investigational Product Adherence	Safety Set
Listing 16.2.6.1	Hamilton Rating Scale for Depression (HAM-D)	Full Analysis Set
Listing 16.2.6.2	Clinical Global Impression (CGI) – Improvement	Full Analysis Set
Listing 16.2.6.3	Clinical Global Impression (CGI) – Severity	Full Analysis Set
Listing 16.2.6.4	Hamilton Anxiety Rating Scale (HAM-A)	Full Analysis Set
Listing 16.2.6.5	Montgomery-Asberg Depression Rating Scale (MADRS)	Full Analysis Set
Listing 16.2.6.6	Short-Form 36 (SF-36) Individual Item Scores	Full Analysis Set
Listing 16.2.6.7	Short-Form 36 (SF-36) Domain/Component Scores	Full Analysis Set
Listing 16.2.6.8	Patient Health Questionnaire (PHQ-9)	Full Analysis Set
Listing 16.2.6.9	Hamilton Rating Scale for Depression (HAM-D) with Percent Retention (%), Response, Remission, Relapse and Rebound	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 Investigational Product	Safety Set
Listing 16.2.7.5	Adverse Events leading to Premature Discontinuation of Investigational Product	Safety Set
Listing 16.2.7.6	Adverse Events leading to Premature Withdrawal from the Study	Safety Set
Listing 16.2.7.7	Adverse Events leading to Dose Reduction	Safety Set
Listing 16.2.8.1.1	Serum Chemistry Data (excluding Liver Function Tests)	Safety Set

Listing Number	Title	Analysis 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	Liver Function Tests	Safety Set
Listing 16.2.8.1.2.1	Potentially Clinically Significant Values of 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 16.2.8.2.1	Potentially Clinically Significant Vital Signs Values	Safety Set
Listing 16.2.8.3	Electrocardiogram (ECG) Data	Safety Set
Listing 16.2.8.3.1	Potentially Clinically Significant QTcF Values	Safety Set
Listing 16.2.8.4	Physical Examination	Safety Set
Listing 16.2.8.5.1	Columbia-Suicide Severity Rating Scale (C-SSRS) – Suicidal Ideation	Safety Set
Listing 16.2.8.5.2	Columbia-Suicide Severity Rating Scale (C-SSRS) – Suicidal Behavior	Safety Set
Listing 16.2.8.6	Physician Withdrawal Checklist (PWC)-20	Safety Set