Protocol Title:	*A Phase 2, Multicenter, Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy and Safety of NV-5138 in Adults with Treatment Resistant Depression*
Protocol Number:	*NAV-17A-007*
Protocol Version	*5.0 (Amendment 4.0)*
Protocol Date	*07 June 2024*
Statistical Analysis Plan Version:	*1.0*
Statistical Analysis Plan Date:	*03 July 2024*

Approvals

Document History

Not applicable

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Abbreviation	Definition
AE	Adverse Event
ASA	American Statistical Association
ATC	Anatomical Therapeutic Chemical
BPRS+	Brief Psychiatric Rating Scale – Positive Symptom Subscale
BMI	Body Mass Index
C-SSRS	Columbia Suicide Severity Rating Scale
CADSS	Clinician Administered Dissociative States Scale
CFB	Change from Baseline
CGI-I	Clinical Global Impression – Improvement
CGI-S	Clinical Global Impression – Severity
CI	Confidence Interval
CRF	Case Report Form
CRO	Clinical Research Organization
CSFQ	Changes in Sexual Functioning Questionnaire
DSMB	Data Safety Monitoring Board
ECG	Electrocardiogram
EMA	European Medicines Agency
EOS	End of Study
FAS	Full Analysis Set

Abbreviation	Definition
FDA	Food and Drug Administration
GAD-7	Generalized Anxiety Disorder 7-item
GCP	Good Clinical Practice
HAM-D_6	Hamilton Depression Rating Scale – 6 Item
HR	Heart Rate
ICE	Intercurrent Event
ICH	International Conference on Harmonization
IPD	Important Protocol Deviation
IRT	Interactive Response Technology
MADRS	Montgomery–Åsberg Depression Rating Scale
MAR	Missing at Random
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple imputation
MMRM	Mixed Model for Repeated Measurements
MNAR	Missing Not at Random
PHQ-9	Patient Health Questionnaire 9-item
PK	Pharmacokinetic
PP	Per Protocol
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD	Standard Deviation

Abbreviation	Definition
SDS	Sheehan Disability Scale
SE	Standard Error
SOC	System Organ Class
SOP	Standard Operating Procedure
TEAE	Treatment-Emergent Adverse Event
TRD	Treatment Resistant Depression
WHO-DD	World Health Organization Drug Dictionary

1. Overview

The NAV-17A-007 study has two versions of the protocol for the purpose of masking the fact that all study subjects will receive placebo in Week 5. This statistical analysis plan (SAP) provides details of statistical methods and reporting for the unmasked Protocol NAV-17A-007 version 5.0, dated 07 June 2024. This unmasked SAP is developed by the Sponsor, Supernus. Both the unmasked protocol and SAP are kept confidential from the clinical research organization (CRO), [redacted], and all study site personnel who are conducting the study until the database is locked.

Unknown to the unmasked protocol, [redacted] develops an SAP and SAS programs for generating tables, figures and listings based on the masked protocol. Thus, after the database lock, all inferential statistics, including p-values, based on a Mixed Model for Repeated Measurements (MMRM) will be generated by Supernus using the data from Baseline and Weeks 1-4 only, without Week 5. All other analysis results generated by [redacted] will remain as is.

2. Study Objectives and Endpoints

2.1. Study Objectives

2.1.1. Primary Objective

The primary objective is to evaluate the efficacy of NV-5138 compared to placebo when administered to adults with Treatment Resistant Depression (TRD) on the Montgomery–Åsberg Depression Rating Scale (MADRS) total score.

2.1.2. Secondary Objectives

2.1.2.1. Key Secondary Objective

The key secondary objective is to evaluate the efficacy of NV-5138 compared to placebo in adult subjects with TRD on the clinician's impression of the severity of depressive symptoms.

2.1.2.2. Additional Secondary Objectives

Additional secondary objectives are to evaluate the efficacy of NV-5138 compared to placebo in adults with TRD on the following:

1) Clinical response to Hamilton Depression Rating Scale-6 Item (HAM-D_6) scale
2) Onset of clinical response
3) Depressive symptoms response
4) Depressive symptoms remission
5) Clinician's impression of improvement of depressive symptoms
6) Individual disability
7) Anxiety Symptoms

8) Clinical response rate of severity
9) Clinical response rate of improvement

[redacted]

2.1.4. Safety Objective

The safety objective is to evaluate the safety and tolerability of NV-5138 compared to placebo in adults with TRD.

2.1.5. Exploratory Objectives

The exploratory objectives are to evaluate the efficacy of NV-5138 compared to placebo in adults with TRD on the following:

1) Self-reported depressive symptoms
2) Self-reported cognitive and physical function
3) Self-reported sexual function

An additional exploratory objective is Pharmacogenomic evaluations.

2.2. Estimands

2.2.1. Primary Estimand

The attributes below are used to construct the primary estimand, defining the treatment effect of NV-5138.

1. Treatment: To compare the treatment effect of NV-5138 versus placebo.
2. Population: The population targeted for the scientific question is defined by the inclusion and exclusion criteria in adults with TRD.
3. Variable: Change from Baseline (Day 1) in MADRS total score to the end of treatment period (Week 4).
4. Intercurrent event: The treatment policy strategy will be followed for handling intercurrent events (ICEs). For subjects who are discontinued due to AE(s) or lack of efficacy and for subjects who are dosed but have no post-baseline MADRS total score, missing data will be imputed using multiple imputation (MI) under the Missing Not at Random (MNAR) assumption; no imputation will be done for subjects discontinued due to other reasons.
5. Population-level summary: The difference in mean change in MADRS total score between NV-5138 and placebo groups will be compared using the mixed model for repeated measures (MMRM) methodology.

2.2.2. Secondary Estimand

The attributes below will be used to construct the secondary estimand regarding the key secondary endpoint, defining the treatment effect of NV-5138.

1) Treatment: To compare the treatment effect of NV-5138 versus Placebo.
2) Population: The population targeted for the scientific question is defined by the inclusion and exclusion criteria in adults with TRD.
3) Variable: Change from Baseline (Day 1) in the Clinical Global Impression – Severity (CGI-S) score to the end of treatment period (Week 4).
4) Intercurrent event: The treatment policy strategy will be used for handling ICEs. For subjects who are discontinued due to AE(s) or lack of efficacy and for subjects who are dosed and have no post-baseline CGI-S scores, missing data will be imputed using MI under the MNAR assumption; no imputation will be done for subjects discontinued due to other reasons.
5) Population-level summary: The difference in mean change in CGI-S scores between the NV-5138 and placebo groups will be compared using the MMRM methodology.

2.3. Study Endpoints

2.3.1. Efficacy Endpoints

2.3.1.1. Primary Efficacy Endpoint

The primary efficacy endpoint of this study is the change from Baseline (CFB) to the end of treatment period (Week 4) on the MADRS total score.

2.3.1.2. Secondary Efficacy Endpoint(s)

The key secondary efficacy endpoint of this study is the CFB to the end of treatment period (Week 4) on the CGI-S score.

Additional secondary efficacy endpoints include the following:

1) CFB to the end of each scheduled week on HAM-D_6 total score
2) CFB to the end of each scheduled week on the MADRS total score
3) Proportion of responders (defined as ≥50% reduction from Baseline in MADRS total score) at the end of each scheduled week
4) Proportion of subjects in remission (MADRS total score ≤10) at the end of each scheduled week
5) Clinical Global Impression – Improvement (CGI-I) score at the end of each scheduled week
6) CFB to the end of each scheduled week on the Sheehan Disability Scale (SDS) total score
7) CFB to the end of each scheduled week on the Generalized Anxiety Disorder 7-item (GAD-7) scale total score
8) Percentage of subjects with a CGI-S score of 1 or 2 at the end of each scheduled week
9) Percentage of subjects with a CGI-I score of 1 or 2 at the end of each scheduled week

2.3.3. Safety Endpoints

The safety endpoints of this study include the following:

1) Adverse events (AEs)
2) Clinical safety laboratory test results
3) Vital signs (including orthostatic blood pressure/pulse rate)
4) Body Weight
5) Electrocardiograms (ECGs)
6) Physical examination
7) Suicidal ideation and behavior as measured by the Columbia Suicide Severity Rating Scale (C-SSRS) score
8) Dissociative symptomatology as measured by the Clinician Administered Dissociative States Scale (CADSS) total score
9) Psychopathology severity as measured by the Brief Psychiatric Rating Scale - positive symptom subscale (BPRS+) score

2.3.4. Exploratory Endpoints

The exploratory endpoints of this study include the following:

- CFB to the end of each scheduled week on the following:
  1) Patient Health Questionnaire 9-item (PHQ-9) scale total score
  2) Massachusetts General Hospital Cognitive and Physical Functioning Questionnaire (CPFQ) total score
  3) Changes in Sexual Functioning Questionnaire (CSFQ) total score

- Pharmacogenomics Endpoints
  1) CFB in the concentration of brain-derived neurotrophic factor (BDNF) will be analyzed as a Pharmacogenomic endpoint
  2) Correlation between baseline pharmacogenetic testing results and drug effects on the efficacy, safety, and/or PK of NV-5138 as an exploratory pharmacodynamic endpoint

3. Overall Study Design and Plan

3.1. Overall Design

This is a multicenter, randomized, double-blind, flexible-dose, placebo-controlled, parallel design of adjunctive NV-5138 in adults with TRD. Approximately 800 outpatient subjects will be screened into the study and 268 subjects will be randomized to NV-5138 or placebo in a 1:1

ratio. Following up to 6 weeks of Screening, there will be a 5-week double-blind treatment period (4 weeks of planned study medication [SM] and 1 week of placebo) and a safety follow-up phone call occurring 30 days after the last dose of blinded study medication. The total study duration from the Screening visit to the end of the treatment period is up to 15 weeks.

3.2. Sample Size and Power

Two hundred subjects (100 per treatment group) will yield 80% power in detecting superiority of NV-5138 to placebo at a significance level of 0.05 (two-sided) using a two-sample t-test with equal allocation to the treatment groups. This assumes an effect size of 0.40 (e.g., mean difference of 4.8 with the standard deviation of 12.0). Assuming a dropout rate of 25%, approximately 268 subjects will be randomized.

Assuming a 67% rate of screen failure, approximately 800 will be screened to obtain approximately 268 subjects randomized.

3.3. Study Population

The study population are adults (aged 18 to 70 years old) with a confirmed diagnosis of TRD.

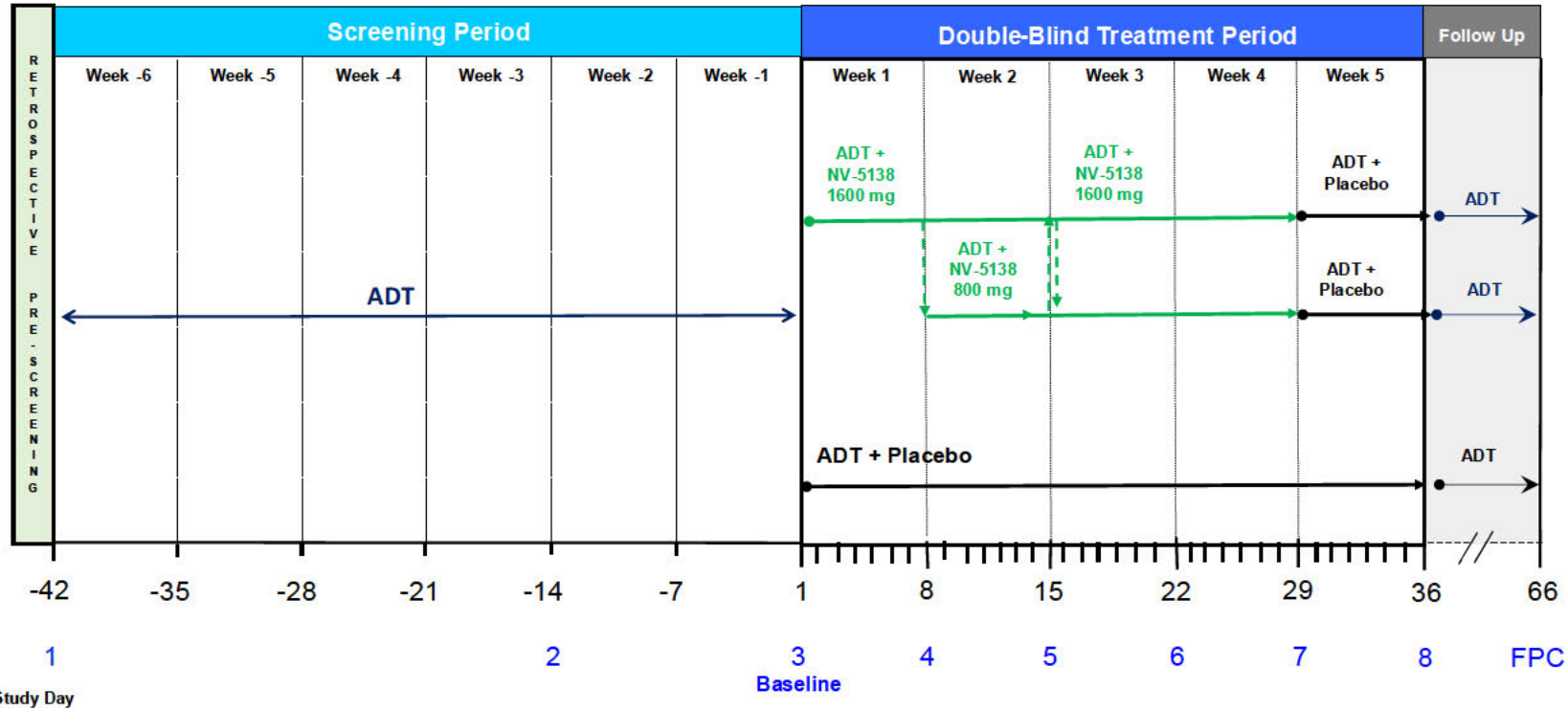
3.4. Treatments Administered

Subjects were allocated to receive one of the following medications at baseline and to take the medication once daily:

- NV-5138, initially at 1600 mg but the dose can be tapered down to 800 mg based on subject's clinical response and tolerability
- Matching placebo

The dosing regime is outlined in Figure 1.

Figure 1 Study Schedule



FPC = Follow Up Phone Call

1. Visit 1 may range from Day -42 to -28 for a subject depending on the length of time in the screening period.
2. Dosing is flexible: the starting dose of NV-5138 or matching placebo is 1600 mg once daily. Dose can be adjusted per Investigator judgment, to maximi ze their treatment response at Visit 4 (Day 8) and Visit 5 (Day 15). Subjects who experience an intolerable adverse effect at 1600 mg may have their dose reduced to 800 mg at Visit 5; subjects who have an inadequate response to 800 mg may have their dose increased again to 1600 mg at Visit 5.
3. Weeks 1-4:Treatment Period, Week 5: Placebo-Washout Week

3.5. Method of Assigning Subjects to Treatment Groups

Allocation of study treatment will occur centrally via an interactive response technology (IRT) using a randomization schedule in a 1:1 ratio to determine the SM assignment for each subject being randomized.

3.6. Blinding and Unblinding

The study subject and all personnel involved with the conduct and interpretation of the study, including the Investigators, study site personnel, the Sponsor and contract research organization (CRO) clinical staff, including the Medical Monitor, will be blinded to study medication. A limited number of Supernus personnel will perform the bioanalytical PK assays and will be aware of the plasma concentration during the clinical study conduct. These personnel should not have access to the randomization schedule, nor be associated with the clinical conduct of the study and will not reveal to any clinical personnel involved in the study the treatment to which a subject is assigned. Randomization schedule data will be kept strictly confidential, filed securely by the IRT vendor, and accessible only to authorized persons until the time of unblinding.

Planned Unblinding

The study blind will be broken only after the study database is locked.

Emergency Unblinding. Emergency unblinding will be performed only under the circumstances outlined in the protocol. The Investigator will promptly document and explain to the Sponsor any premature unblinding (e.g., accidental unblinding, unblinding due to a serious adverse event) of treatment assignment.

3.7. Schedule of Events

The schedule of events is outlined in the following table:

Table 1 Schedule of Events

Period	Screening		Double-Blind Treatment						Safety Follow -up*
Visit Number	**1**	**2**	**3 (Baseline)**	**4**	**5**	**6**	**7**	**8 EOS/ET**	
Day of Study	**-42[a]**	**-14**	**1**	**8**	**15**	**22**	**29**	**36**	**66**
Week of Study	**-6**	**-2**	**–**	**1**	**2**	**3**	**4**	**5**	
Study Visit Window (Days)	**–**	**± 1**	**± 2**	**± 2**	**± 2**	**± 2**	**± 2**	**± 2**	**± 7**
Signed informed consent (ICF) including PGx consent	√								
Relevant histories (social, medical, psychiatric, family psychiatric, neurological)	√								
Mini-International Neuropsychiatric Interview (MINI)	√								
[redacted]									
ATRQ Investigator Rated	√								
Demographics and Height	√								
Inclusion/Exclusion criteria	√	√	√						
Alcohol consumption	√		√	√	√	√	√	√	
Recreational cannabis use[c]	√	√	√	√	√	√	√	√	
Physical examination	√		√					√	
Vital signs[d] and Weight	√	√	√	√	√	√	√	√	
ECG	√		√	√	√	√	√	√	
Blood sample for FSH (post-menopausal females only)	√								
Blood sample for Hematology & Chemistry	√		√	√	√	√	√	√	
[redacted]									
Blood sample for PGx[f]			√[f]			√[f]			
Blood sample for ADT	√	√	√	√	√	√	√	√	
Serum pregnancy test (FOCP only)	√								

Period	Screening		Double-Blind Treatment						Safety Follow -up*
Visit Number	**1**	**2**	**3 (Baseline)**	**4**	**5**	**6**	**7**	**8 EOS/ET**	
Day of Study	**-42[a]**	**-14**	**1**	**8**	**15**	**22**	**29**	**36**	**66**
Week of Study	**-6**	**-2**	**–**	**1**	**2**	**3**	**4**	**5**	
Study Visit Window (Days)	**–**	**± 1**	**± 2**	**± 2**	**± 2**	**± 2**	**± 2**	**± 2**	**± 7**
Urine pregnancy test (FOCP only)			√	√	√	√	√	√	
Urine sample for urinalysis	√		√	√	√	√	√	√	
Urine drug and alcohol screen[g]	√	√	√	√	√	√	√	√	
Randomization			√						
Subject PRR training video		√			√				
Register medication Diary	√								
Review medication adherence[h]		√[h]	√[h]	√[h]	√[h]	√[h]	√[h]	√[h]	
MADRS	√	√	√	√	√	√	√	√	
CGI-S	√	√	√	√	√	√	√	√	
CGI-I				√	√	√	√	√	
HAM-D_6			√	√	√	√	√	√	
SDS			√			√	√	√	
GAD-7			√			√	√	√	
CADSS	√		√	√	√	√	√	√	√
C-SSRS[i]	√	√	√	√	√	√	√	√	√
BPRS+	√		√	√	√	√	√	√	√
CPFQ			√	√	√	√	√	√	
PHQ-9			√			√	√	√	
CSFQ			√			√	√	√	
Review adverse events			√	√	√	√	√	√	√
Review concomitant medication	√	√	√	√	√	√	√	√	√
SM dispensed			√	√	√	√	√		
SM returned and accountability				√	√	√	√	√	
Follow-up phone call									√

ADT = antidepressant therapy (ADT); ATRQ = Antidepressant Treatment Response Questionnaire; BPRS+ = Brief Psychiatric Rating Scale

positive symptom subscale; CADSS = Clinician Administered Dissociative States Scale; CGI-I = Clinical Global Impression – Improvement scale; CGI-S = Clinical Global Impression – Severity scale; CPFQ = Massachusetts General Hospital Cognitive and Physical Functioning Questionnaire; CSFQ = Changes in Sexual Functioning Scale; C-SSRS = Columbia Suicide Severity Rating Scale; ECG = electrocardiogram; EOS = end of study; EOS = end of study; ET=early termination; FOCP = females of childbearing potential; FSH = follicle stimulating hormone; GAD-7 = Generalized Anxiety Disorder 7-item Scale; HAM-D_6 = Hamilton Depression Rating Scale (6 items); ICF = informed consent form; MADRS = Montgomery-Asberg Depression Rating Scale; MINI=Mini-International Neuropsychiatric Interview; PGx=pharmacogenomics; PHQ-9 = Patient Health Questionnaire-9 Item; PK = pharmacokinetics; POC= Point of Care; PRR = placebo response reduction; [redacted] SDS = Sheehan Disability Scale; SM = study medication

* The Safety Follow-up phone call will occur approximately 30 days after the last dose of blinded SM only for subjects who completed the study. For subjects who withdraw from the study, a safety follow-up phone call will occur approximately 7 days after withdrawal.

a. Day of the Study for Visit 1 will be depending on the prescreening ADT history (Day -42, Day -35, or Day -28)
[redacted]
c. Recreational use of cannabis is allowed (not daily, and not exceeding 3 times per week for either vape, smoke, or edible), and should be kept consistent during the study. Subjects must agree to refrain from using cannabis 48 h prior to the study visits.
d. Orthostatic blood pressure, pulse rate, respiratory rate, and oral temperature.
[redacted]
f. Samples will be stored for analysis and will not be used for identification. At Visit 3 two samples will be collected: one for pharmacogenomic (DNA) and one for BDNF analysis. At Visit 6, one sample will be collected for BDNF analysis.
g. Perform “Standard UDS and alcohol from Visit 1 and Visit 2. At Baseline (Visit 3) and all post-Baseline visits POC test will be performed for UDS, while breathalyzers will be provided for alcohol screen.
h. Sites should review the use of the medication Diary at each visit and intervene with participants as needed.
i. Screening/Baseline C-SSRS version will be assessed at Visit 1 and the “since last visit” version will be assessed thereafter.

4. Statistical Methods

The analyses of treatment comparisons described in this SAP, with the exception of outputs identified as necessary for the Drug Safety Monitoring Board (DSMB) meetings, will not be performed until the study database is locked and treatment is unblinded.

Outputs for the DSMB meetings will be produced by an independent biostatistician governed by the DSMB charter.

4.1. General Principles

All statistical analyses will be performed using SAS (release 9.4 or higher). Supernus will generate all inferential statistics, including p-values, based on a Mixed Model for Repeated Measurements (MMRM) using the data from Baseline and Weeks 1-4 only, without Week 5. All other statistical results will be generated by [redacted], which is the designated CRO. [redacted] will be responsible for creating TLF reports and delivering the reports as well as the programs to Supernus at the completion of the study.

Continuous (quantitative) variable summaries will include the number of subjects (n) with non-missing values, mean, standard deviation (SD), median, interquartile range (Q1 and Q3), minimum, and maximum.

Categorical (qualitative) variable summaries will include the frequency and percentage of subjects who are in the particular category or each possible value. In general, the denominator for the percentage calculation will be based upon the total number of subjects in the study population, unless otherwise specified. The denominator for by-visit displays will be the number of subjects in the relevant study population with non-missing data at each visit.

The minimum and maximum will be reported with the same degree of precision (i.e., the same number of decimal places) as the observed data. Measures of location (mean, median, Q1 and Q3) will be reported to 1 degree of precision more than the observed data and measures of spread (SD) will be reported to 2 degrees of precision more than the observed data.

Percentages will be presented to 1 decimal place, unless otherwise specified.

Unless otherwise indicated, all statistical tests will be conducted at the 5% significance level using 2-tailed tests, and *P* values will be reported accordingly. If specifically specified, corresponding 95% two-sided confidence interval (CI) will be presented.

Baseline is defined as the last observation measured at or before the first dose of study treatment.

4.2. Interim Analysis and Data Monitoring

No formal interim efficacy analyses are planned. A DSMB has been established to review safety data. Details are described in the DSMB charter.

4.3. Analysis Population

The following analysis populations are planned for this study:

- **Randomized Population:** The Randomized population includes all subjects who are randomized.
- **Full Analysis Set (FAS):** The FAS is a subset of subjects in the Randomized Population who take at least one dose of SM and have a Baseline assessment of MADRS total score. The efficacy analyses will be conducted using the FAS according to the randomized treatment assignment.
- **Per Protocol (PP) Population:** The PP Population is a subset of subjects in the FAS who complete all 8 visits through EOS with no missing MADRS total score assessments and no important protocol deviations (Section 5.2). Subjects in the PP Population will be analyzed according to the actual treatment received.
- **Safety Population:** The Safety Population includes all subjects randomized into the study who receive at least one dose of SM. The safety analysis will be performed according to the actual treatment received.

[redacted]

Assignment of subjects to the Per Protocol populations will be confirmed at a blinded data review meeting to be held before the study database is locked.

4.4. Adjustments for Covariates

The efficacy analyses will include covariates in the MMRM model (Section 6.1.1). For all other analyses, unless otherwise specified, no adjustment for covariates will be used.

4.5. Multiple Comparisons

A hierarchical hypothesis testing structure will be used to ensure the overall alpha is 0.05. Formal hypothesis testing on the key secondary endpoint will only be performed if the null hypothesis is rejected for the primary endpoint.

4.6. Handling of Dropouts or Missing Data

With respect to the primary and key secondary analysis, the MMRM method, implemented via SAS® PROC MIXED (SAS/STAT Software), will be used for handling missing MADRS total scores under the Missing at Random (MAR) assumption. However, the missing data for subjects who are discontinued due to AE or lack of efficacy, and for subjects who are dosed but have no post-baseline MADRS total score, will be assumed missing not at random (MNAR). Placebo-based multiple imputation will be used to fill in missing data for subjects with data assumed to be MNAR. This approach assumes that missing data from the NV-5138 treatment group would adopt the outcomes estimated from the placebo group. The SAS procedure, PROC MI, will be used to apply the MNAR statement with the option of using the observations from the placebo group (see Section 6.1.2 for PROC MI code that can be used to perform imputations for subjects with data assumed to be MNAR). Multiple imputation will be carried out by the fully conditional specification (FCS) method. One hundred (100) imputed datasets will be created for those subjects who discontinued due to AE or lack of efficacy or who are dosed but have no post-baseline MADRS total score. The 100 sets of the imputed data for each of these subjects will be

used to replace the missing data before performing the MMRM analysis for the primary endpoint. Finally, to combine the 100 MMRM results, PROC MIANALYZE in SAS will be used to make a final statistical inference.

The sensitivity analysis for the primary and key secondary endpoint will be performed by assuming missing MADRS total scores are MNAR using placebo-based multiple imputation. The same procedure will be applied to key secondary endpoint CGI-S.

For analysis of other secondary endpoints, missing values will be handled in the same fashion as the primary efficacy endpoint.

For safety analyses, missing dates for AEs and concomitant medications will be imputed as described in Section 4.10 of this SAP. Missing data for all other safety endpoints will not be imputed.

4.7. Analysis Visit Windows

Analyses of all variables for this study will be based upon the visit at which the data was collected. Data from unscheduled visits will be excluded from all summaries by visit, but data from unscheduled visits will be in the data listings. The efficacy data collected at the ET visit will be mapped to the next scheduled visit per protocol.

4.8. Pooling of Sites

This is a multicenter study. Small sites (defined as those with sample size < 10 FAS subjects/site) will be pooled. All the sites will be sorted by the number of FAS subjects. The smallest site with less than 10 FAS subjects will be combined with the second smallest site. Then all the sites will be sorted again till there is no more site with less than 10 FAS subjects. The MMRM will be performed including pooled site as a factor.

4.9. Derived Variables

Details of derived variables are given below.

1) The MADRS is a 10-item Investigator-rated diagnostic questionnaire used to measure the severity of depressive episodes in subjects with mood disorders and is designed to be sensitive to changes brought on by the treatment. Each question is scored from 0 to 6, the sum of the 10 subtests score will yield a total score ranging from 0 to 60. If one item score is missing, then the value for the missing items will be imputed using the mean of the non-missing items rounded to the nearest integer. If more than one item score is missing, the total score will not be calculated. A total score ranging from:
    a. 0-6 = normal range (no depression)
    b. 7-19 = mild depression
    c. 20-34 = moderate depression
    d. ≥35 = severe depression
2) The CGI-S is evaluated on a 7-point scale as outlined below:
    a. 1 = Normal, not at all depressed
    b. 2 = Borderline depressed

c. 3 = Mildly depressed
d. 4 = Moderately depressed
e. 5 = Markedly depressed
f. 6 = Severely depressed
g. 7 = Among the most extremely depressed patients

3) The Hamilton depression rating scale is one of the most widely used clinician-administered depression scales. The original version contains 17 items related to symptoms of depression over the past week developed for hospital inpatients. The HAM-D_6 (6-items version), derived from the original 17-items version of the scale, offers sensitivity for measuring severity of detecting improvement of depression comparable to other more complex versions. Five of the six items (Depressed Mood, Work and Activities, Somatic Symptoms General, Feelings of Guilt, Anxiety Psychic, Retardation) are scored on a scale of 0 to 4, and one item (Somatic Symptoms General) is scored on a scale of 0 to 2, for a total score of 0-22. If any item score is missing, the total score will not be calculated. A total score ranging from:
   a. 0-4 = normal range (no depression)
   b. 5-8 = mild depression
   c. 9-12 = moderate depression
   d. 13-22 = severe depression
4) The CGI-I is evaluated by the Investigator at each post-Baseline study visit during treatment relative to the subject's condition at Baseline on a 7-point scale as outlined below:
   a. 1 = Very much improved
   b. 2 = Much improved
   c. 3 = Minimally improved
   d. 4 = No change
   e. 5 = Minimally worse
   f. 6 = Much worse
   g. 7 = Very much worse
5) The SDS is a short self-report tool that assesses functional impairment in three interrelated domains consisting of 1) Work/School, 2) Social Life and 3) Family Life/Home Responsibilities. Each of the items is scored on an 11-point scale (0-10) on which 0 indicates no impairment, 1-3 mild impairment, 4 to 6 moderate impairment, 7-9 marked impairment and 10 extreme severe impairment. Scores from each item are summed to measure global functional impairment ranging from 0 (unimpaired) to 30 (highly impaired). Scores exceeding 5 points on any of the items are indicative of functional impairment.
   For the Work/School domain, if the box is checked, then the depressive symptoms on "work/school" domain are not applicable. In this case, the total SDS score will be derived from the other two scales ("social life" and "family life/home responsibilities") and then multiplied by a ratio of 3/2 (so, ranging from 0 to 30). If the box isn't checked and the score of the scale on "work/school" is missing, then the total SDS score will be considered missing. If the score for Social Life or Family Life/Home Responsibility domain is missing, the total score will be missing.

6) The Generalized Anxiety Disorder 7 scale (GAD-7) is a self-reported 7-item questionnaire for Screening and measuring the severity of generalized anxiety disorder. The subject scores each GAD-7 item as outlined below:
    a. 0 = Not at All
    b. 1 = Several Days
    c. 2 = More than half the days
    d. 3 = Nearly Every Day

   The total score is obtained by summating all 7 items scores. The total score ranges from 0 to 21:
    a. 0-4 = none/minimal anxiety
    b. 5-9 = mild anxiety
    c. 10-14 = moderate anxiety
    d. ≥ 15 = severe anxiety

   If one item score is missing, then the value for the missing items will be imputed using the mean of the non-missing items rounded to the nearest integer. If more than one item score is missing, the total score will not be calculated.
7) The Cognitive and Physical Functioning Questionnaire (CPFQ) is a subject rated scale which measures cognitive and physical functionality. Subjects rate themselves on seven common aspects of cognitive function on a six-point scale, where 1=greater than normal, 2 = normal, 3 = minimally diminished, 4 = moderately diminished, 5 = markedly diminished and 6 = totally absent. The total score ranges from 6 to 42, and is categorized as follows:
    a. ≤7 = Greater than normal functioning
    b. 8–14 = Normal functioning
    c. 15–21 = Minimally diminished functioning
    d. 22–28 = Moderately diminished functioning
    e. 29–35 = Markedly diminished functioning
    f. 36–42 = Totally absent functioning

   If one item score is missing, then the value for the missing items will be imputed using the mean of the non-missing items rounded to the nearest integer. If more than one item score is missing, the total score will not be calculated.
8) The CSFQ is a subject rated scale which measures sexual function. Subjects rate themselves on 14 items using a 5-point scale, and the total score reflects the level of sexual dysfunction. The total score ranging from 14–70 is the sum of the 14 items. A score of ≤47 for male subjects, and ≤41 for female subjects, indicates sexual dysfunction. If one item score is missing, then the value for the missing items will be imputed using the mean of the non-missing items rounded to the nearest integer. If more than one item score is missing, the total score will not be calculated.
9) The PHQ-9 is a subject rated scale that measures overall subject health. Subjects are asked to rate how often a set of nine situations have concerned them over the last 2 weeks, giving a total score that ranges from 0 to 27 where:
    a. 0 = not at all
    b. 1 = several days
    c. 2 = more than half of the days
    d. 3 = nearly every day

If one item score is missing, then the value for the missing items will be imputed using the mean of the non-missing items rounded to the nearest integer. If more than one item score is missing, the total score will not be calculated.

The following safety variables will also be derived.

1) The CADSS is a 23-item clinician administered scale that measures present-state dissociative symptoms. The subject then rates items on a 4-point Likert scale ranging from:
    a. 0 = not at all
    b. 1 = mild
    c. 2 = moderate
    d. 3 = severe
    e. 4 = extreme

    The total score ranges from 0 to 92.
    If one item score is missing, then the value for the missing items will be imputed using the mean of the non-missing items rounded to the nearest integer. If more than one item score is missing, the total score will not be calculated.
2) The C-SSRS is an assessment tool that evaluates suicidal ideation and behavior. It consists of 10 categories, to which a yes or no response is required. The following endpoints are derived from the tool:
    a. Self-Injurious behavior without Suicidal Intent
    b. Suicidal ideation - A Yes response to one of questions 1 to 5
    c. Suicidal behavior - A Yes response to one of questions 6 to 10
    d. Suicidal ideation or behavior – A Yes response to any of questions 1 to 1
3) The Brief Psychiatric Rating Scale (BPRS+) is a tool clinicians or researchers uses to measure psychiatric symptoms such as anxiety, depression, and psychoses. For the study, the version used is adapted from the original BPRS 4-item positive symptoms (BPRS+). The scale is comprised of 4 items assessing a subject's experience of psychosis, often referred as positive symptoms of Suspiciousness, Hallucinatory Behavior, Unusual Thought Content and Conceptual Disorganization. Of the 4 items assessed, the first three items are based on questions asked by the clinician to the subject, Conceptual Disorganization is a clinician rated item based on observation of subjects' behavior and speech during the assessment. Each subscale is rated on a 8-point scale ranging from 0-7 as follows:
    a. 0 = not able to be assessed
    b. 1= not present
    c. 2 = very mild
    d. 3 = mild
    e. 4 = moderate
    f. 5 = moderately severe
    g. 6 = severe
    h. 7 = extremely severe

    The total score ranging from 0-28 is the sum of the 4 items.

4.10. Data Handling/Conventions

All collected data will be presented in listings.

All *P* values will be displayed in four decimals and rounded using standard scientific notation (e.g., 0.XXXX). If a *P* value less than 0.0001 occurs, it will be shown in tables as < 0.0001.

Adverse events will be coded using the MedDRA version 24.0.

A treatment related AE is any AE with a relationship to the SM of possibly or definitely related.

For Adverse Events and Concomitant Medications, the following rules will be applied for partial dates:

If just the day is missing then the day assigned is the first day of the month or the date of first dose (if in the same month), whichever is later; if just the month is missing then the month assigned is the month of the first dose, unless that results in a date before the first dose in which case the month after the first dose is used; and if both the month and the day are missing then the month assigned is the month of the first dose and the day assigned is either the first day of the month or the first dose date, whichever is later.

If partial times occur, the convention is as follows: if the missing time occurs on the day of the first dose and both the hour and minute are missing then the time assigned is the time of the first dose, otherwise if both the hour and minute are missing and the date is not the date of first dose the time assigned is 12:00; if the date is the same as the date of the first dose and only the hour is missing the hour assigned is 12 or the hour of first dose, whichever is later, and if the date is the same as the date of first dose and only the minute is missing the minute assigned is 30 or the minute of first dose, whichever is later. Otherwise the hour assigned is 12 if the hour is missing and the date is not the same as the date of first dose and the minute assigned is 30 if the the date is not the same as the date of first dose.

5. Study Subjects and Demographics

5.1. Subject Disposition

Disposition will include tabulations of the number of subjects randomized into each treatment group, the number of subjects who received treatment, tabulated reasons for discontinuation from the study, and the number of subjects in each analysis population.

5.2. Important Protocol Deviations

A protocol deviation is defined as any change, divergence, or departure from the study design, or procedures defined in the protocol. An important protocol deviation (IPD) is defined as a protocol deviation that might significantly affect the completeness, accuracy, and/or reliability of the study data or that might significantly affect a subject's rights, safety, or well-being.

Potential IPDs may include, but are not limited to:

- Subjects enrolled into the study even though they did not meet all the inclusion criteria or met any of the exclusion criteria
- Use of any prohibited concomitant medication after randomization
- Subject is dispensed or took incorrect dose of SM

Occurrence of any of these and other events may be considered as potential IPDs. However, for the analysis these events should be categorized as IPDs only if they could potentially affect interpretation of study results. Prior to database lock and break of treatment blind, potential IPDs identified in the clinical database and/or recorded in the site deviation log will be reviewed for the determination of IPDs, which would lead to the exclusion from the PP Population. IPDs will be provided in subject data listings.

5.3. Demographics and Other Baseline Characteristics

Summary statistics for age, gender, race including race group (White vs Non-White), ethnicity, height, weight, and Body Mass Index (BMI) will be presented by treatment group and overall. The baseline disease status, as shown by the baseline MADRS total score, current ADT, anxiety, prior ADT response, duration of current ADT, duration of the major depressive episode, cannabis use, speech latency, CGI-S and HAM-D6 total scores, and the number of ADTs with inadequate response will also be presented.

For the continuous variables, the number of non-missing values and the mean, standard deviation, minimum, median, interquartile range (Q1 and Q3), and maximum will be tabulated.

For the categorical variables, the counts and proportions of each value will be tabulated.

These analyses will be conducted for the Safety Population and FAS.

The number and percent of subjects reporting various medical histories (Social, Medical, Psychiatric, Family psychiatric and Neurological), grouped by MedDRA system organ class and preferred term (coded using MedDRA v. 24.0), will be tabulated by treatment group. This analysis will be conducted for the Safety Population.

5.4. Exposure and Compliance

Duration of exposure is defined as the total number of days a subject is exposed to any study medication. This will be calculated for each subject by taking the difference between the date of last dose minus the date of the first dose, plus 1 (date of last dose minus date of first dose +1). Duration of treatment exposure will be summarized by using descriptive statistics (n, mean, SD, median, interquartile range (Q1 and Q3), minimum, and maximum).

Duration of treatment exposure will be categorized as follows: 1-7 days, 8-14 days, 15-21 days, 22-28 days, 29-35 days and 35 days or more.

Subject exposure will be summarized categorically by study week, dose (800 mg, 1600 mg in separate rows: i.e., 1600 mg for week 1, 800 or 1600 mg for rest weeks), and treatment group. Study weeks are defined as days 1-7, 8-14, 15-21, 22-28 and 29-35. The number and percent of subjects in each dose group will be presented for each treatment.

Percentage compliance will be calculated as the following:

$$\left(\frac{\sum_{i=3}^{7}(\#of\ capsules\ dispensed\ at\ visit_i - \#of\ capsules\ returned\ at\ visit_{i+1})}{(\sum_{i=3}^{7}(Y_i\ \times \#\ of\ days\ from\ visit_i\ to\ visit_{i+1}))}\right) \times 100$$

where Y_i equals the number of capsules that the subject was instructed to take daily between $visit_i$ and $visit_{i+1}$.

For each treatment, percentage compliance will be summarized by compliance category (<80%, 80-120%, and >120%) and number of subjects in each compliance category. Study medication compliance will also be summarized as a continuous variable using descriptive statistics (n, mean, SD, median, interquartile range [Q1 and Q3], minimum, and maximum) for each treatment.

Exposure and compliance summaries will be performed for the Safety population.

6. Efficacy Analysis

All efficacy analyses, unless otherwise stated, will be performed using the FAS. Week 4 is the real end of active treatment period. Thus, Week 5 will be excluded from analyses based on MMRM methodology. All efficacy data including Week 5 will be summarized using descriptive statistics and presented in data listings.

6.1.1. Primary Efficacy Analysis

The change from Baseline in MADRS total score to the end of treatment period (Week 4), will be analyzed using a MMRM. For subjects who discontinued the study because of Adverse Event (AE) or lack of efficacy, and for subjects who are dosed but have no post-baseline MADRS total score, missing data will be imputed based on the MNAR assumptions as described in Section 4.6. For data missing for any other reason, no imputation will take place. The data collected at the ET visit will be mapped to the next scheduled visit.

The MMRM model will include Baseline MADRS total score as a covariate, and pooled site, treatment, visit, and treatment-by-visit interaction as fixed effects. The model parameters will be estimated using the restricted maximum likelihood method with an unstructured variance-covariance matrix and Kenward-Roger approximation to estimate the denominator degrees of freedom.

If the MMRM model with the unstructured variance-covariance matrix does not converge using the default Newton-Raphson fitting algorithm by PROC MIXED in SAS, the Fisher scoring algorithm or the no-diagonal factor analytic structure will be used (Lu & Mehrotra, 2010).

If the model still fails to converge, the following types of covariance structure with the sandwich estimators (using the EMPIRICAL option) will be used to fit the model in a sequential order until the model converges:

1. Heterogeneous Toeplitz
2. Heterogeneous Autoregressive of order 1

3. Toeplitz
4. Autoregressive of order 1
5. Compound symmetry

The denominator degrees of freedom will be estimated using the BETWITHIN option in SAS when any of the covariance structures is used with the sandwich variance estimator. The Least Squares (LS) Mean of change from Baseline to end of treatment period in the MADRS total score for each treatment group will be presented, along with the corresponding standard error. The NV-5138 group will be compared with the placebo group by presenting the difference in the LS Means between NV-5138 and placebo groups (NV-5138 minus placebo) with its 95% confidence interval and the p-value.

The mean profile over time for the MADRS total score will be plotted by treatment group.

Descriptive statistics (n, mean, SD, median, interquartile range (Q1 and Q3), minimum, and maximum) will be tabulated.

6.1.2. Sensitivity Analysis

The sensitivity analysis will be performed using the placebo-based multiple imputation to fill in missing MADRS total scores. This approach assumes that after discontinuation, subjects from the NV-5138 treatment group would adopt the outcome model estimated from the placebo arm. The placebo-based imputation will be implemented by adopting the following three steps.

1. SAS PROC MI (SEED=220877) is applied to the input dataset containing all by-visit MADRS Total Score during the baseline and treatment period. Multivariate imputation will be carried out by the fully conditional specification (FCS) method. One hundred (100) multiply-imputed datasets will be created. SAS® PROC MI with the MNAR statement will be implemented by using the following SAS code.

```
PROC MI SEED=220877 NIMPUTE=100;
 CLASS treatment;
 VAR treatment baseline week1 week2 week3 week4;
 FCS REG;
 MNAR MODEL (baseline week1 week2 week3 week4 /MODELOBS=(treatment=("1"));
RUN;
```

In the above SAS code, baseline, week1, week 2, week 3, and week 4 are analysis values of MADRS total scores at baseline and at each of the treatment weeks from week 1 to week 4, where MODELOBS=(treatment=("1"))) means that observations will be from the placebo arm only.

Imputed values will be rounded to the closest integer. In case imputed values are below or above the range of the MADRS scale, values will be imputed to respectively the minimum or maximum value of the scale.

2. For each of the imputations from Step #1, Change from Baseline (CFB) to study visit will be computed on the 100 completed datasets.

For each imputation, the CFB will be analyzed using the MMRM model in the same manner as the primary efficacy analysis (Section 6.1.1).

3. Finally, PROC MIANALYZE will be used to make the final inference based on the 100 results of the MMRM model from Step 2.

The adjusted mean (LS Mean) of CFB to End of Study (EOS) of NV-5138 treatment group and placebo, along with the corresponding standard error, difference in the LS Means between the NV-5138 and placebo groups (NV-5138 minus placebo), 95% confidence intervals for the treatment difference and the p-value will be presented from the combined estimates.

6.2. Secondary Efficacy Analyses

The secondary analyses will be based on the FAS as stated in Section 4.3. Unless otherwise stated, all secondary analyses will use MMRM based on the same model described for the primary analysis on the change from Baseline to the end of treatment period except for CGI-I, which analyzes the absolute value with the Baseline CGI-S as a covariate. The NV-5138 treatment arm will be compared with the placebo arm. The LS means of the treatment groups, differences in the LS Means between the NV-5138 and placebo groups (NV-5138 minus placebo), and 95% CIs for the treatment differences with p-values will be computed.

Where the use of the MMRM model is not appropriate or not used, an appropriate alternative will be used based on the individual endpoints as outlined in the following sections.

For responder analyses, the Chi-squared Test will be used.

6.2.1. Key Secondary Efficacy Analyses

The change from baseline to the end of treatment period in the CGI-S score will be analyzed in the same way as the primary analysis (Section 6.1.1). Specifically, intercurrent events, missing data, and statistical method for the key secondary efficacy endpoint analysis will be handled as described in Sections 2.2 and 6.1. Formal hypothesis testing on the key secondary endpoint will only be performed if the null hypothesis is rejected for the primary endpoint. Sensitivity analysis will be conducted in the same manner as described in Section 6.1.2.

6.2.2. Analyses of Additional Secondary Endpoints

The following secondary endpoints will be analyzed and presented:

1) The change from baseline to the end of each scheduled week on HAM-D_6 total score will be analyzed using MMRM as described for the primary analysis.
2) The change from baseline in MADRS total score to the end of each scheduled week will be analyzed using MMRM as described for the primary analysis.
3) The proportion responders (defined as ≥50% reduction from Baseline in the MADRS total score) at the end of each scheduled week will be presented for each treatment group. The 2-sided 95% CI around the difference in proportions (NV-5138 minus placebo) and the p-value from Chi-squared Test will be presented.

4) The proportion of subjects in remission (defined as MADRS total score ≤10) at the end of each scheduled week will be analyzed as in item 3 above.
5) The absolute value of CGI-I score will be analyzed at the end of each scheduled week in the same way as the primary analysis using the baseline CGI-S score as a covariate in the MMRM model.
6) The change from baseline to the end of each scheduled week in the SDS will be analyzed using MMRM as described for the primary analysis.
7) The change from baseline to the end of each scheduled week in the GAD-7 will be analyzed using MMRM as described for the primary analysis.
8) CGI-S will be dichotomized at each visit as 'responder' if the score is 1 or 2 and 'non-responder' if the score is 3 to 7. Percentage of subjects with CGI-S score of 1 or 2 at the end of each scheduled week will be summarized using categorical analysis in the same method as described in item 3 above.
9) CGI-I will be dichotomized at each visit as 'responder' if the score is 1 or 2 and 'non-responder' if the score is 3 to 7. Percentage of subjects with CGI-I score of 1 or 2 at the end of each scheduled week will be summarized using categorical analysis in the same method as described in item 3 above.

6.3. Exploratory Efficacy Analysis

6.3.1. PHQ-9

For PHQ-9 scale, summary statistics of the total score will be produced for each visit. The change from baseline at the end of each scheduled week in the total score will be analyzed using MMRM as described for the primary analysis on the FAS in Section 6.1.1.

6.3.2. CPFQ and CSFQ

For the CPFQ and CSFQ scales, summary statistics of the total scores will be produced for each visit. The change from baseline at the end of each scheduled week in the total score will be analyzed using MMRM as described for the primary analysis on the FAS in Section 6.1.1.

6.4. Supplementary Analysis

The primary and key secondary analyses will be repeated using the PP Population.

6.5. Subgroup Analysis

The MADRS total score, CGI-S, HAM-D_6 and SDS will be repeated for the following subgroups:

- Age group (18-40; 41-60; >60 years)
- Gender (Male; Female; Non-Binary)
- Race (White; Non-White)

- Baseline MADRS total score (split by median: ≤ median; > median)
- Current ADT (Type of ADT – defined by WHO Drug generic term)
- Anxiety (anxiety; no anxiety). Anxiety will be identified by the MINI (subjects are counted as having anxiety if they have social anxiety or generalized anxiety)
- Inadequate ADT response for the current MDE [redacted] (≥2 to 4 ADTs with ≤25% response; ≥2 to 4 ADT with 26% to <50% response only; 1 ADT with ≤25% response; 1 ADT with 26% to <50% response only)
- Duration of current ADT (split by median: ≤ median; > median)
- Duration of the major depressive episode (MDE) (split by median: ≤ median; > median)
- Cannabis use (using urine drug test result): Positive at baseline; Negative at baseline and positive at least one post-baseline Visit; Negative at baseline and all post-baseline visits
- Speech Latency (abnormal vs normal): A single split Recursive Decision Tree (RDT) will be used to determine the "cutoff" latency value which will be calculated from post-randomization data (MADRS total score ≥ 20 vs MADRS total score ≤ 10). The latency cutoff will be applied to the pre-randomization sessions to classify subjects into normal or abnormal.

7. Safety and Tolerability Analysis

Safety analyses will be performed by treatment arm based on the Safety Population.

7.1. Adverse Events

All adverse events after the first dose will be considered Treatment Emergent Adverse Events (TEAE).

All TEAEs will be coded using the MedDRA Coding Dictionary, Version 24.0.

The number and percent of subjects reporting TEAEs, grouped by MedDRA system organ class (SOC) and preferred term (PT), will be tabulated in the following groups:

- All TEAEs
- TEAEs by severity
- TEAEs by relationship to study medication
- SAE
- SAE related to study medication
- AE leading to death
- TEAE leading to study medication withdrawn
- TEAE leading to study medication dose modification (medication interrupted, dose reduced, dose increased)

In the case of multiple occurrences of the same TEAE within the same subject, each subject will only be counted once for each preferred term.

In the summaries showing severity and relationship to study medication the event with the maximum severity or strongest relationship will be reported. If a particular event is missing the

severity and/or relationship, then the strongest possible severity or relationship will be assumed for analysis (severity = severe, relationship = definitely related).

A summary of incidence rates (frequencies and percentages) of TEAEs by treatment group, SOC, and preferred term will be prepared for the Safety Population. Common TEAEs (≥ 5%) by treatment group and preferred term will be prepared.

In the AE data listings, all AEs will be displayed.

7.1.1. Treatment Emergent Adverse Events Leading to Study Medication Withdrawn

A summary of incidence rates (frequencies and percentages) of TEAEs leading to study medication withdrawn, by treatment group, SOC, and preferred term will be prepared for the Safety Population. No inferential statistical tests will be performed.

A data listing of AEs leading to withdrawal of study drug will also be provided, displaying details of the event(s) captured on the CRF.

7.1.2. Treatment Emergent Adverse Events Leading to Study Medication Dose Modification

A summary of incidence rates (frequencies and percentages) of TEAEs leading to study medication dose modification, by treatment group, SOC, and preferred term will be prepared for the Safety Population. No inferential statistical tests will be performed.

A data listing of AEs leading to dose modification of study drug will also be provided, displaying details of the event(s) captured on the CRF.

7.1.3. Deaths and Serious Adverse Events

Any deaths that occur during the study will be listed.

Serious adverse events will be listed and tabulated by system organ class and preferred term and presented by treatment group.

7.2. Clinical Laboratory Evaluations

Laboratory data presentations will use SI units.

Laboratory test (Hematology and Chemistry) results will be summarized descriptively by treatment and time point as both observed values and change from baseline values. For continuous variables, summary statistics (n, mean, SD, median, Q1, Q3, minimum and maximum) will be provided. For categorical variables, the number and percentage of subjects in each category will be reported.

The number of subjects with clinical laboratory values below, within, or above the normal range, by time point, and in relation to baseline will be tabulated for each clinical laboratory analyte by treatment group (shift table).

Urinalysis results will be summarized at each time point. For parameters with a numerical result, the mean value at each visit will be presented, as well as the shift relative to the normal range.

A listing of "Abnormal" or "Out of Range" will be provided.

7.3. Vital Signs

Vital signs will be collected at all study visits. Descriptive summaries of actual values and changes from baseline will be summarized for weight (kg), BMI (kg/m2), oral body temperature (ºC), respiration rate (breaths per minute), sitting pulse rate (bpm), standing pulse rate (bpm), sitting systolic blood pressure (mmHg), standing systolic blood pressure (mmHg), sitting diastolic blood pressure (mmHg), standing diastolic blood pressure (mmHg). The change in blood pressure and pulse rate from sitting to standing will be summarized. A subgroup analysis of CFB in weight by BMI (<30 vs ≥30) will be summarized. These summaries will be presented by study visit and treatment group.

The number of subjects with vital signs values below, within, or above normal ranges, by study visit will be tabulated (shift tables) for each vital sign parameter by treatment group. Normal ranges are presented in Table 2.

A listing of "Abnormal" or "Out of Range" will be provided.

Table 2 Vital Sign Normal Ranges

Measurement	**Normal Range**
Body Mass Index	18.5 – 24.9 kg/m^2
Temperature	95-100 ºF (35-37.8 ºC)
Diastolic blood pressure	60 – 90 mmHg
Systolic blood pressure	90 – 140 mmHg
Pulse rate	50 – 100 bpm
Respiration rate	10 – 25 breaths per minute

7.4. Electrocardiograms

Descriptive summaries will be presented for ECG measures of PR interval, QRS interval, QT interval, QTcF and Heart Rate (HR). These summaries will be presented by visit and treatment group.

The number and percentage of subjects with normal and abnormal ECG results will be summarized for the Safety Population by treatment group for each visit.

Additionally, the number and percentage of subjects with QTcF values ≥ 450 msec (male) and ≥ 470 (female) and with changes from baseline <= 30 msec, 30 <- ≤60 msec, and >60 msec will be presented by treatment group for each visit and the worst post-baseline result.

A listing of "Abnormal" or "Out of Range" will be provided.

7.5. Further Safety Evaluations

Physical Examination reports which are clinically significant will be listed.

C-SSRS outcomes will be summarized using number and percent of subjects by categories for suicidal ideation only, suicidal behavior only, suicidality (ideation and behavior combined) and non-suicidal self-injurious behavior. The summary will be presented by treatment group. The number of subjects with/without suicidal ideation only, suicidal behavior only, suicidality (ideation and behavior combined), by study visit will be tabulated (shift tables) by treatment group.

Descriptive statistics will be presented for CADSS total score, and BPRS+ total score by visit and treatment group.

7.6. Prior and Concomitant Medications

Prior and concomitant medications will be summarized descriptively by treatment group using counts and percentages. Medications will be coded using WHO-DD dated March 2021.

Prior medications will be presented separately from concomitant medications. Medications that started before first dose of SM will be considered prior medications whether or not they were stopped before Baseline. Any medications continuing or starting after first dose of SM will be considered to be concomitant. If a medication starts before Baseline and continues after Baseline it will be considered both prior and concomitant.

7.7. Antidepressants (ADTs)

The following are the antidepressants approved for the study: citalopram, escitalopram, paroxetine, fluoxetine, sertraline, duloxetine, venlafaxine (IR or XR), desvenlafaxine, vilazodone, levomilnacipran, vortioxetine, bupropion and bupropion/dextromethorphan. Antidepressants will be summarized by treatment using counts and percentages. Medications will be coded using WHO-DD dated March 2021.

8. PK and Pharmacogenomic Analysis

[redacted]

8.2. BDNF Analysis

CFB in BDNF will be analyzed descriptively.

8.3. DNA Analysis

Samples for DNA will be stored and may be analyzed after DBL. The analysis will not be covered in this SAP.

9. Changes from the Protocol

Pooled site is added in the primary efficacy MMRM model. Hierarchical hypothesis testing framework added to allow formal hypothesis testing for the key secondary endpoint.

10. Validation

All tables, listings and figures will undergo a full QC process, in line with current [redacted] prior to release to the sponsors. The documentation for this process will be stored in the eTMF.

11. References

ASA. (2018) Ethical Guidelines for Statistical Practice. Prepared by the Committee on Professional Ethics, April 2018. http://www.amstat.org/about/ethicalguidelines.cfm

ICH (1998). ICH Harmonized Tripartite Guideline. Statistical Principles for Clinical Trials E9; 1998. https://database.ich.org/sites/default/files/E9_Guideline.pdf

ICH (2019). ICH Harmonized Tripartite Guideline. Addendum on Estimands and Sensitivity Analysis in Clinical Trials to the Guideline on Statistical Principles for Clinical Trials E9(R1); https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf

Lu, K and Mehrotra, D. Specification of covariance structure in longitudinal data analysis for randomized clinical trials. Statistics in Medicine 2010:29;474-488

RSS. (2014) The Royal Statistical Society: Code of Conduct, 2014.https://rss.org.uk/RSS/media/File-library/About/2019/RSS-Code-of-Conduct-2014.pdf.