STATISTICAL ANALYSIS PLAN

STUDY TITLE:

A Phase 2, Randomized, Double-Blind, Placebo-Controlled, Parallel-Group Study to Evaluate the Efficacy, Safety, and Tolerability of TNX-601 ER Monotherapy versus Placebo in Patients with Major Depressive Disorder (MDD) "UPLIFT"

PROTOCOL NUMBER:

TNX-TI-M201

SHORT TITLE:	UPLIFT
NCT#:	NCT05686408
COMPOUND #:	
CLIENT:	Tonix Pharmaceuticals
IND NUMBER(S):	[redacted]
PREPARED BY:	[redacted]

[redacted]

ACKNOWLEDGEMENT AND SIGNATURE SHEET

[redacted]

[redacted]

VERSION HISTORY

SAP Version	Version Date	Change(s)	Rationale
[redacted]			

[redacted]

TABLE OF CONTENTS

[redacted]

[redacted]

LIST OF TABLES

[redacted]

1. LIST OF ABBREVIATIONS

Table 1: List of Abbreviations

Abbreviation	Term
AE	adverse event
AESI	adverse event of special interest
BID	twice daily
BMI	body mass index
CFR	Code of Federal Regulations
CGI-I	Clinician Global Impression of Improvement
CGI-S	Clinician Global Impression of Severity
CI	confidence interval
CMH	Cochran-Mantel-Haenszel
CRF	case report form
CSFQ-14	Changes in Sexual Functioning Questionnaire Short-Form
CSR	clinical study report
C-SSRS	Columbia Suicide Severity Rating Scale
ECG	electrocardiogram
EMA	European Medicines Agency
ET	early termination
FDA	Food and Drug Administration
GCP	Good Clinical Practice
ICH	International Council on Harmonization of Technical Requirements for Pharmaceuticals for Human Use
IND	Investigational New Drug
IRB	Institutional Review Board
IP	Investigational product
IU	international unit
IWRS	interactive web response system
LOE	lack of efficacy
LS, LSM	least square, least square means
mITT	modified intent-to-treat
MedDRA	Medical Dictionary for Regulatory Activities
MAR	missing at random

[redacted]

MCMC	Markov Chain Monte Carlo
MI	multiple imputation
MMRM	Mixed-Effects Model of Repeated Measures
NDA	New Drug Application
NSAIDs	Non-steroidal anti-inflammatory drugs
PGI-C	Patient Global Impression of Change
PGI-S	Patient Global Impression of Severity
QD	once daily
SAE	serious adverse event
SAP	statistical analysis plan
SD	standard deviation
SF	screen failure
SOC	system organ class
TEAE	Treatment-Emergent Adverse Event
US	United States
WHO-DD	World Health Organization Drug Dictionary

2. PURPOSE OF THE ANALYSES

This statistical analysis plan (SAP) describes the planned analysis and reporting for Protocol TNX-TI-M201 (A Phase 2, Randomized, Double-Blind, Placebo-Controlled, Parallel-Group Study to Evaluate the Efficacy, Safety, and Tolerability of TNX-601 ER Monotherapy versus Placebo in Patients with Major Depressive Disorder).

The structure and content of this SAP provides sufficient detail to meet the requirements identified by the US Food and Drug Administration (FDA), European Medicines Agency (EMA), and International Council for Harmonisation (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use: Guidance on Statistical Principles in Clinical Trials. All work planned and reported for this SAP will follow internationally accepted guidelines for statistical practice, as published by the American Statistical Association and the Royal Statistical Society.

The planned analyses identified in this SAP may be included in clinical study reports (CSRs), regulatory submissions, or future manuscripts. Also, post-hoc exploratory analyses not necessarily identified in this SAP may be performed to further examine study data. Any post-hoc, or unplanned exploratory analysis performed will be clearly identified as such in the final CSR.

[redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

The reader of this SAP is encouraged to also read the clinical protocol, and other identified documents, for details on the planned conduct of this study. Operational aspects relating to collection and timing of planned clinical assessments are not repeated in this SAP unless they are relevant to the planned analysis.

[redacted]

3. PROTOCOL SUMMARY

This is a 6-week, randomized, double-blind, placebo-controlled, parallel-group, multicenter study to evaluate the efficacy, safety, and tolerability of TNX-601 ER, 39.4 mg monotherapy, on symptoms of depression in adult patients with MDD in approximately 30 sites in the United States. This study plans to randomize approximately 300 male and female patients aged 18 to 65 years old (inclusive), with current MDD as defined by DSM-5 criteria. During the 6-week double-blind treatment period, patients will receive either TNX-601 ER or placebo in a 1:1 randomization ratio. After the End of Treatment (Visit 5/ET), there will be a 2-week posttreatment safety follow-up period.

The study will be conducted in 4 periods: Screening (Visit 1), Baseline (Visit 2), Treatment (Visits 3-5), and Post Study Follow-up (Visit 6).

Screening (Visit 1)

Patients will be consented and assessed for current MDD as defined by DSM-5 diagnostic criteria, using the Mini International Neuropsychiatric Interview, Version 7.0.2 (MINI 7.0.2). Depression severity will be assessed by MADRS. Eligibility will also depend on medical history, physical and neurological examination, use of concomitant medications, and the results of laboratory tests (eg, pregnancy and drug screen, liver, kidney, and thyroid function, electrocardiogram).

If a patient meets all eligibility requirements but is currently taking a prohibited medication, there will be a Screening Period of up to 35 days for drug tapering and washout. For extenuating circumstances and in cases in which a prospective participant is on fluoxetine, the duration of the Screening Period may be increased to up to 49 days with Medical Monitor approval. Medication-free washout time for prohibited medications will be 4 weeks for antidepressants (except for fluoxetine, which will be 6 weeks) and atypical antipsychotics.

Baseline, Randomization, Start of Treatment (Visit 2)

To ensure continued eligibility prior to randomization, patients will be assessed for all inclusion/exclusion criteria, including a MADRS total score of ≥ 25. Patient must also have a $\leq$25% decrease or increase in MADRS score from Screening (Visit 1) to Baseline (Visit 2). Any patients not meeting these randomization criteria will be considered a Screen Failure (SF) and will not continue in the study. Eligible patients will be randomized in a 1:1 ratio to one of two treatment arms (once daily oral TNX-601 ER or placebo). At this visit, study drug will be dispensed, and the first dose will be given to eligible patients.

Double-Blind Treatment Period (Visits 3-5)

The Treatment Period will last for about 42 days. Several investigator- and patient-rated scales and safety assessments (eg, C-SSRS, AEs) will be administered at each visit. At each of the 3 visits, study drug compliance and accountability will be assessed by collecting previously dispensed study drug bottle and counting the remaining tablets.

[redacted]

All patients who withdraw early from the study will be encouraged to complete an Early Termination (ET) visit as early as possible for safety evaluations. The patients will also be strongly encouraged to come back for the Post Study Follow-up (Visit 6).

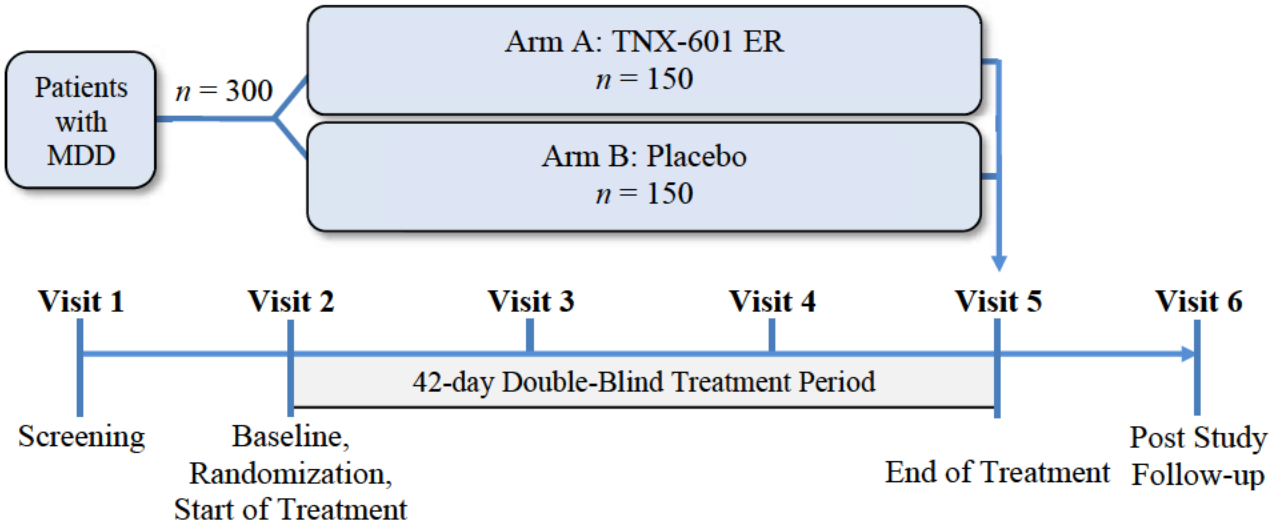
Post Study Follow-up (Visit 6)

Between discontinuation of study drug and the Post Study Follow-up (Visit 6), potential withdrawal effects will be assessed at home using the SOWS at +1 day, +2 days, +3 days, +6 days, +9 days, and +12 days after the End of Treatment (Visit 5/ET). This includes patients who are early terminated.

Approximately 14 days (+7 days) after discontinuation of study drug, all randomized patients will return to their respective study site for an off-drug, Post Study Follow-up (Visit 6). This visit will include an assessment of prior and concomitant medications, vital signs and weight, pregnancy test, AEs, C-SSRS, and the SOWS.

The overall duration in the study for each patient is expected to be up to 13 weeks.

Figure 1: Study Schema



The study timeline and events schedule are provided in Figure 1.

[redacted]

4. ANALYSIS SAMPLES

Three populations will be utilized for the analyses of data. Disposition of patients will use the All Patients Population. The safety analyses will be conducted on the Safety Population. The intention-to-treat (ITT) Population will be used for the efficacy analyses.

4.1. All Patients Population

This includes all patients available in the electronic data capture system; patients will be presented by their randomized treatments (if applicable).

4.2. Intention-to-Treat Population

The ITT population will include all randomized patients who receive at least 1 dose of the investigational product. The ITT population is the primary population for efficacy analyses and patients will be analyzed based on their randomized treatment.

4.3. Safety Population

The safety population will include all patients who received at least one dose of study drug. This population will be used to summarize the safety data. Patients will be analyzed according to their treatment received; if an error in issuing investigational product (IP) occurs and a patient receives a mix of treatments, they will be analyzed under the treatment arm with the most IP issued to the patient.

5. ESTIMANDS FOR PRIMARY AND KEY SECONDARY EFFICACY ENDPOINTS

5.1. Primary Efficacy

Population

The target population is individuals with major depressive disorder as defined by the inclusion/exclusion criteria.

Variable

The primary endpoint will be the change from Baseline (Visit 2) to Week 6 in the MADRS total score.

Intercurrent Events

Concomitant medications and all other intercurrent events other than discontinuation will use a treatment policy strategy; data will be analyzed as observed. A hypothetical strategy will be employed to handle intermittent missing data and discontinuation. Patients that that discontinue for adverse events or lack of efficacy will be assumed to be representative of patients that would not continue treatment and their values will be imputed with multiple imputation (MI) using the placebo distribution. All other missing values, whether intermittent or following discontinuation will be assumed to be missing at random and imputed with MI within treatment group. See section 9.1.1 for details of the imputation.

Population-Level Summary

The population-level summary will be the difference in least-square means between treatment arms (analyzed as randomized) at Week 6.

The primary efficacy analysis will use a mixed model repeated measures (MMRM) approach to estimate mean change from baseline to Week 6 in the MADRS total score visit in the TNX-601 ER and placebo arms. The model will include all patients in the ITT population, and the dependent variable will be the observed change from baseline in the MADRS total score at each post-randomization in-clinic visit. Covariates in the model will include the fixed categorical effects of treatment, site, visit, and treatment by visit interaction, as well as the continuous fixed covariates of baseline score and baseline score by visit interaction. Full details are given in section 9.1.2.

5.2. Key Secondary Efficacy

The efficacy endpoints listed below are considered key secondary endpoints.

- Change from Baseline (Visit 2) in the CGI-S score at Week 6
- Change from Baseline (Visit 2) in the SDS total score at Week 6

The key secondary efficacy endpoints will apply estimands parallel with the primary estimand, adjusted for the endpoint in question, but with identical populations and handling of intercurrent

[redacted]

events. MMRM methodology will be utilized for the continuous secondary endpoints, and analyses will be based on the ITT population. The dependent variable will be the observed change from baseline in the respective secondary endpoint at each post-randomization visit. Covariates in the model will include the fixed categorical effects of treatment, site, visit, and treatment by visit interaction, as well as the continuous fixed covariates of baseline score and baseline score by visit interaction.

[redacted]

6. STUDY PATIENTS

6.1. Disposition of Patients

The number and percentage of patients in each analysis population will be summarized. The number of patients screened, patients enrolled, and the number and percentage of patients completing the study, discontinuing from the study and reasons for discontinuation will also be reported. All percentages will be calculated using the number of patients enrolled as the denominator.

6.2. Demographic and Other Baseline Characteristics

Demographic and baseline characteristics such as age, gender, race/ethnicity, height, weight, BMI, and smoking history will be summarized by treatment group (TNX-601 ER and placebo) and overall for the ITT and Safety populations. Categorical items will be reported with counts and percentages; continuous items will be reported using summary statistics (mean, standard deviation (SD), median, minimum, maximum).

6.3. Prior and Concomitant Medications/ Therapies

Prior medications are defined as medications or therapies initiated prior to the start of the study drug and terminated prior to the start of study drug. Hence, these medications or therapies will have end dates prior to the first dose date of study drug. Concomitant medications are defined as any medications other than the study drug that a patient receives concurrently with the study drug. These medications will have end dates on or after the first dose date of the study drug if they have an end date. In the case of missing or partial dates, medications will be considered concomitant unless the available data excludes that possibility.

All medications will be coded using the World Health Organization Drug Dictionary (WHO-DD). Prior and concomitant medications will be summarized by treatment group and by the number and percentage of patients taking each medication, classified by using WHO-DD Anatomical Therapeutic Chemical Level 3 and preferred term.

Prior medications for MDD are collected on a separate case report form including whether the patient had an efficacy response to the drug and how well they tolerated the drug. Prior MDD drugs will be summarized by preferred term and response; and by preferred term and tolerability.

Prior psychotherapy will be collected at the screening visit. The count and percentage of patients receiving each type of therapy will be summarized by treatment for each prior therapy; this will also be reported for each therapy flagged as ongoing. All psychotherapy will be presented in a listing including the reported start and stop dates of the therapy.

All medications and therapies will be presented in listings, with prior MDD medications and therapies listed separately.

6.4. Medical History and Physical Examination

Medical History will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) and summarized by system organ class (SOC) and Preferred Term using frequency counts and percentages by treatment group. Physical examination findings will be presented in listings.

[redacted]

7. STUDY OPERATIONS

7.1. Protocol Deviations

A protocol deviation is any noncompliance with the clinical trial protocol or ICH Good Clinical Practice (GCP) requirements. The noncompliance may be either on the part of the patient, the investigator, or the study site staff. As a result of deviations, corrective actions are to be developed by the site and implemented promptly. Protocol deviations will be logged and categorized as important or non-important prior to unblinding. Protocol deviations will be summarized by type, status as important/non-important, and by treatment group for the Safety population.

Individual protocol deviations will be listed by patient.

7.2. Randomization

This is a double-blind study; unless otherwise specified, all study personnel are to remain blinded to study drug prior to database lock and final analysis following study unblinding.

A total of approximately 300 patients will be randomized in a 1:1 ratio to treatment with once-daily oral TNX-601 ER (tianeptine hemioxalate extended-release tablets) or identical placebo tablets, which are devoid of the tianeptine hemioxalate but contain the same inactive ingredients present in the active tablets. Randomization will be stratified by study site.

The randomization schedule will be computer generated using a permuted block algorithm and will randomly allocate study drug to randomization numbers. The randomization numbers will be assigned through a central interactive web response system (IWRS) as patients are entered into the study.

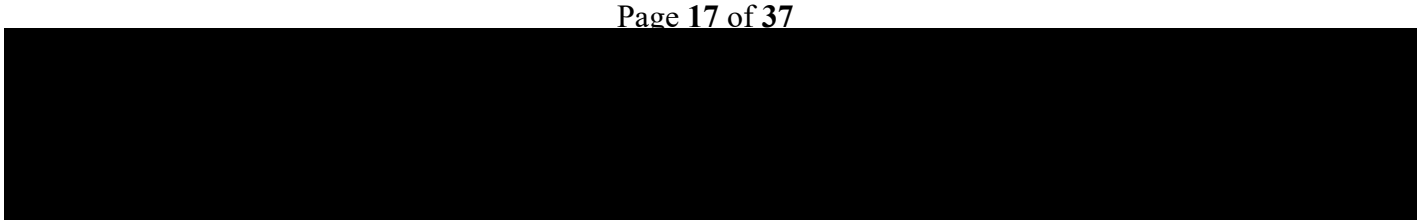
7.3. Measures of Treatment Compliance

The treatment duration will be calculated and summarized based on CRF data for first and last dose dates (number of days=last dose date – first dose date+1).

Days of exposure will be based on drug accountability and the number of tablets taken assuming once daily dosing. The number of patients with total exposure by visit weeks (≤2 weeks, 2 to ≤4, 4 to ≤6, and >6 weeks) will be presented.

Compliance will be similarly summarized across all study visits and overall for each treatment arm. Study drug compliance as a percentage will be defined as the exposure days defined above divided by the total number of expected days on treatment multiplied by 100. The expected number of days will be the drug return date-drug issue date +1 for individual visits and the treatment duration as defined above for the overall compliance.

Compliance will be summarized with descriptive statistics by treatment arm. The number and percentages of patients within certain categories of compliance (e.g. < 60%, 60% to < 80%, 80% to ≤ 100%, greater than 100%) will also be presented.



A listing of drug accountability data based on CRF data will be provided.

8. GENERAL ANAYLSIS AND REPORTING CONSIDERATIONS

Descriptive summaries will be provided where appropriate for each of the primary and secondary variables. In general, tables will summarize data by treatment group and visit.

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

Categorical, qualitative variable summaries will include the frequency and percentage of patients who are in the particular category. In general, the denominator for the percentage calculation will be based upon the total number of patients in the study population for the treatment group unless otherwise specified.

Unless otherwise specified, baseline values are defined as the last non-missing measurement prior to the first dose of study drug. Change from baseline will be defined as the post-baseline visit value minus the baseline value.

For continuous endpoints where treatment groups are compared via linear models with the difference in LS means, the effect size will be calculated and included as a summary statistic using the difference in LS means and its standard error, transformed to a standard deviation. Study day is defined as assessment date – first dose date +1 for dates after the first dose date and assessment date – first dose date for dates prior to the first dose.

All analyses will be performed using SAS® Software version 9.4 or later.

8.1.1. Study Success

[redacted]

8.1.2. Multicenter Studies

This is a multicenter study. Unless otherwise specified, data from all study sites will be pooled for all analyses unless otherwise specified.

8.1.3. Assessment Time Windows

Patients who withdraw/drop out from the study will have the early termination (ET) data collected at their ET visit included in the analysis at the closest visit where a given assessment should be collected per the schedule of assessments (Week 2, 4, 6), using midpoints between

[redacted]

visits to window the early termination. If this results in two records for a given visit, then the scheduled visit will be used, followed by the record closest to the targeted date. See section 10.5.4.3 for details specific to the SOWs windowing.

8.1.4. Timing of Analyses

All final, planned analyses will be performed after the last participant has completed all study assessments, all relevant study data have been processed and integrated into the analysis database, and the database has been locked.

Any post-hoc, exploratory analyses completed to support planned study analysis, which were not identified in this SAP, will be documented and reported in appendices to the CSR. Any results from these unplanned analyses (post-hoc) will also be clearly identified in the text of the CSR.

8.1.5. Multiple Comparisons/Multiplicity

The overall type I error rate will be controlled at α=0.05 for the set of primary and key secondary endpoint comparisons using a serial gatekeeping approach.

At each step, if a p-value fails to meet the alpha threshold for significance, the process will be halted and all subsequent comparisons declared non-significant. This will continue with the same pattern through the endpoints in the order listed below:

The order of the analyses will be:

- Change from Baseline (Visit 2) in the MADRS total score at Week 6 (Primary)
- Change from Baseline (Visit 2) in the CGI-S score at Week 6
- Change from Baseline (Visit 2) in the SDS total score at Week 6

Although all p-values will be non-significant following the first to fail to meet its alpha threshold, p-values will be reported for all the above endpoints; they will be noted as descriptive for those following the first non-significant p-value.

8.1.6. Power and Sample Size

The sample size will be approximately 300 patients with MDD. [redacted]

[redacted]

[redacted]

9. ENDPOINT EVALUATION

9.1. Primary Endpoint

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

The MADRS is a validated, 10-item, clinician-administered depression scale. It will be completed at Screening (Visit 1), Baseline (Visit 2) and after Weeks 2, 4, and 6 of treatment (Visits 3, 4, and 5, respectively). The 10 items measure the core symptoms and cognitive features of clinical depression. This scale has historically been used in drug treatment trials due to its particular sensitivity to detect treatment effects.

Patients with a Screening (Visit 1) MADRS score of ≥28, ≥25 at Baseline (Visit 2), and a ≤25% decrease or increase in MADRS score from Screening (Visit 1) to Baseline (Visit 2) are eligible to be randomized.

9.1.1. Computation of the Primary Endpoint

Each item has a severity scale from 0 to 6, with higher scores reflecting more severe symptoms. Individual items will be summed to form an overall score (from 0 to 60). It is not anticipated that there will be any missing individual items since the raters will be instructed to complete all the assessments.

A multiple imputation approach will be used to impute data. Missing MADRS overall scores for participants in the ITT population will be imputed via multiple imputation (MI). [redacted]

[redacted]

[redacted]

Continuous outcomes included in the list of key secondary outcomes will be handled as above. Seeds for the SAS procedures are given in appendix section 14.2.

9.1.2. Primary Analysis of the Primary Endpoint

Summary statistics (mean, SD, median, minimum, maximum) for the primary efficacy endpoint will be reported for the change from the Baseline to week 6 for each of the treatment groups. The change from Baseline will be analyzed in an MMRM model for each multiple imputation set with fixed effects of treatment, visit, baseline value, treatment by visit interaction, baseline by visit interaction and pooled site; an unstructured covariance matrix will be used. Sites with less than 5 patients will be pooled into a single site; this covariate may be removed from the model in the case of convergence issues. Least square means (LSMs), standard errors, and the LSM differences (treatment vs placebo) for each of the 20 MI repeats will be combined using SAS procedure MIANALYZE (Rubin, 1976) and 95% CIs and p-values will be reported for each. Denominator degrees of freedom will use the Kenward Rogers approach. Should the model with unstructured variance fail to converge using the default Newton-Raphson approach, the Fisher scores (via the SCORING option of the PROC MIXED statement) will be utilized. Should this also fail to converge, then other approaches as detailed in Lu 2010 will be utilized. Once these approaches have been exhausted, then the following structures will be used in order with the first to converge being selected: heterogeneous Toeplitz, heterogeneous first-order autoregressive, Toeplitz, first order autoregressive and compound symmetric will be attempted in that order; in these cases, the sandwich estimator (Diggle and Kenward 1994) will be used to estimate the standard errors of the fixed effects parameters and the denominator degrees of freedom will utilize the DDFM=BETWITHIN option in SAS®.

9.1.3. Sensitivity Analyses of the Primary Analysis

The following sensitivity analyses may be performed:

9.1.3.1. Missing at Random Multiple Imputation

For this imputation approach, all values will be imputed under the assumption that they are missing at random. A staged approach as described in section 9.1.1 will be used to first impute sporadic missing values with MCMC, then a monotone regression approach for the second stage. However, all patients will have missing values imputed within treatment group under the MAR assumption.

9.1.3.2. Missing Not at Random Multiple Imputation: All patients imputed from the placebo distribution

For this imputation approach, a staged approach as described in section 9.1.1 will be used to first impute sporadic missing values with MCMC, then a monotone regression approach for the second stage. However, all patients will have missing values imputed using the placebo distribution as reference. This will inform the hypothetical possibility that all subjects revert to the placebo distribution regardless of their reason for dropout.

[redacted]

9.1.3.3. Multiple Imputation Tipping Point

A tipping point analysis will be performed using the primary analysis MI datasets (from section 9.1.2) and applying an increasing delta (of MADRS score) to the active group until statistical significance is no longer achieved. The delta will only be applied to imputed values; this will be done in 0.5 increments on the MADRS total score. This will only be performed if the primary analysis is significant. This will test the overall robustness of the primary analysis.

9.1.3.4. Current MADRS Version

The first sixteen subjects were inadvertently assessed under an older version of the MADRS (see Clinical Information Amendment dated 25Sep2023 and referenced in section 2). To assess the impact of this, the primary analysis will be repeated with these 16 subjects excluded.

9.1.4. Secondary Analyses of the Primary Endpoint

In addition to the change from baseline to Week 6 in the MADRS total score, the change from baseline at Week 2 and Week 4 will also be reported out of the primary MMRM model described in section 9.1.2.

Additionally, three responder-type analyses of the MADRS total score will be performed:

- Proportion of patients achieving response, defined as a ≥50% decrease from Baseline (Visit 2) in MADRS total score, assessed at Weeks 2, 4, and 6
- Proportion of patients in remission, defined as a MADRS total score <=10, assessed at Weeks 2,4, and 6
- Proportion of patients achieving sustained response, defined as a ≥50% decrease from Baseline (Visit 2) in MADRS total score at Weeks 4 and 6

For each, the number and percentage of responders (with response for each as described above) and non-responders will be reported, with missing values considered non-responders. 95% CIs of the percentages will be given, as well as the difference between treatment groups in percentages, its 95% Wald CI, and p-value comparing the groups. Groups will be tested using a difference in proportions Z test (equivalent to a Pearson's Chi-Squared). In the case of low cell counts (<5 expected in a given cell), then Fisher's exact test will be used instead.

9.2. Secondary Endpoints

All secondary efficacy analyses will be based on the ITT population.

9.2.1. Key Secondary Endpoints

There will be two key secondary endpoints:

9.2.1.1. Change from Baseline (Visit 2) in the CGI-S score at Week 6

The CGI-S score will be scored 1-7 and treated as a continuous variable. It will be analyzed in a manner identical to the primary endpoint (MMRM with MI using an MNAR approach for patients dropping for AEs and LOE).

[redacted]

Sensitivity analyses will include considering all patient dropouts as missing at random (paralleling the approach in section 9.1.3.1) and considering all patient dropouts as missing not at random (paralleling the approach in section 9.1.3.2).

9.2.1.2. Change from Baseline (Visit 2) in the SDS total score at Week 6

For participants that respond to all three questions (work, social life/leisure and home) on the SDS, their total score will be the sum of the three items. Participants that do not respond to the work/school item because they have not worked or attended school in the recall period for reasons unrelated to their disease state will not have that value imputed; to retain the 0 to 30 scaling for the total score, the participant's other two items will be summed and multiplied by 1.5. This approach reports the disease impact on the applicable domains. Due to the design of the ePRO, participants may both respond to the work/school item and check the box that they have not worked or attended school in the recall period for reasons unrelated to their disease state. In these instances, the checkbox will supersede the response and the 0-10 score will be censored; as above, the participant's other two items will be summed and multiplied by 1.5. Participants that have either of the other items missing or have work/school item missing but do not check the box to indicate that they did not work/attend school will be assigned a missing value.

The SDS total score will be analyzed in a manner identical to the primary endpoint (MMRM with MI using an MNAR approach for patients dropping for AEs and LOE).

Sensitivity analyses will include considering all patient dropouts as missing at random (paralleling the approach in section 9.1.3.1) and considering all patient dropouts as missing not at random (paralleling the approach in section 9.1.3.2).

9.2.2. Exploratory Endpoints

9.2.2.1. Continuous Exploratory Endpoints

Each of the following continuous endpoints will be summarized:

- Change from Baseline (Visit 2) in the CGI-S score at Week 2 and Week 4
- Change from Baseline (Visit 2) in the SDS total score at Week 2 and Week 4
- Change from Baseline (Visit 2) in PGIS score at Weeks 2, 4, and 6
- Change from Baseline (Visit 2) to Week 6 in the CSFQ-14 total score in females and in males (analyzed separately)
- Change from Baseline (Visit 2) in the HARS total score at Weeks 2, 4, and 6
- Change from Baseline (Visit 2) in each of the 3 domains (work/school; social life; home life or family responsibilities) using the SDS at Weeks 2, 4, and 6
- Mean PGIC score at Weeks 2, 4, and 6
- Mean CGI-I score at Weeks 2, 4, and 6

The CGI-S score and SDS total score at Week 2 and Week 4 will be reported out of the same MMRM model used for each of these key secondary endpoints.

[redacted]

The PGIS, HARS total score, and SDS domains (work/school; social life; home life or family responsibilities) will be analyzed using a model identical to the primary efficacy endpoint, but on observed data only, without multiple imputation. Additionally, the primary efficacy and key secondary endpoints will be reported out of an MMRM without imputation.

The CSFQ-14 total score will be summarized with descriptive statistics and compared between groups with an ANCOVA with effects for treatment and baseline value.

The PGIC and the CGI-I will be considered both as responder analyses (see below) and as continuous 1-7 values. For continuous PGIC and the CGI-I, summary statistics (mean, SD, median, minimum, maximum) will be reported for each visit for each of the treatment groups. They will be analyzed in an MMRM model with fixed effects of treatment, visit, treatment by visit interaction, and site; an unstructured covariance matrix will be used. Least square means (LSMs), standard errors, and the LSM differences (each treatment vs placebo) with 95% CIs and p-values will be reported for each visit. Denominator degrees of freedom will use the Kenward Rogers approach. Should the model with unstructured variance fail to converge using the default Newton-Raphson approach, the Fisher scores (via the SCORING option of the PROC MIXED statement) will be utilized. Should this also fail to converge, then other approaches as detailed in Lu 2010 will be utilized. Once these approaches have been exhausted, then the following structures will be used in order with the first to converge being selected: heterogeneous Toeplitz, heterogeneous first-order autoregressive, Toeplitz, first order autoregressive and compound symmetric will be attempted in that order; in these cases, the sandwich estimator (Diggle and Kenward 1994) will be used to estimate the standard errors of the fixed effects parameters and the denominator degrees of freedom will utilize the DDFM=BETWITHIN option in SAS®.

9.2.2.2. Categorical Secondary Endpoints

The following categorical endpoints will be tabulated:

- Proportion of patients with a CGI-I rating of “very much improved” or “much improved” at Weeks 2, 4, and 6
- Proportion of patients with a PGIC rating of “very much improved” or “much improved” at Weeks 2, 4, and 6
- Proportion of patients with a CGI-S rating of “normal, not at all ill” or “borderline mentally ill” at Weeks 2, 4, and 6
- Proportion of patients with a PGIS rating of “not present”, “very mild”, or “mild” at Weeks 2, 4, and 6

For each, at each visit, the number and percentage of patients in each category (without dichotomizing) will be reported as well as the number of missing values. The number and percentage of responders (with response for each as described above) and non-responders will be reported, with missing values considered non-responders. 95% CIs of the percentages will be given, as well as the difference between treatment groups in percentages, its 95% Wald CI, and p-value comparing the groups. Groups will be tested using a difference in proportions Z test

[redacted]

(equivalent to a Pearson's Chi-Squared). In the case of low cell counts (<5 expected in a given cell), then Fisher's exact test will be used instead.

9.2.2.3. Subgroup Analyses

Subgroup analyses will be conducted for the primary and key secondary endpoints. These will report the descriptive statistics within each subgroup for observed data without imputation and (if feasible) fit the primary MMRM model and report the LSmeans & 95% CIs for each treatment and timepoint. Treatment differences at each timepoint will be reported with the LSmean difference and 95% CI, if estimable.

Subgroups will include:

- Age (>40 years old, <=40 Years old)
- Sex at birth (Male/ Female)
- Race (White/Non-white)
- Ethnicity (Hispanic or latino/ Not Hispanic or latino)

10. SAFETY EVALUATION

10.1. Overview of Safety Analysis Methods

All safety analyses will be completed for the Safety Population. All safety data will be tabulated and listed. No formal statistics or imputation will be performed for the safety analysis. Safety and tolerability will be assessed by:

- Incidence of AEs and SAEs throughout the entire duration of the study
- Incidence of AEs associated with abuse liability using the MADDERS®
- Assessment of physical examination findings
- Changes from baseline (Visit 2) in ECG results
- Changes from baseline (Visit 2) in clinical laboratory test results
- Changes from baseline (Visit 2) in vital signs and weight
- Indications of increased suicidal ideation or behavior as assessed by the C-SSRS
- Changes from baseline (Visit 2) in patient-rated CSFQ-14 in males and in females, assessed separately
- Change from baseline Visit 5 or End of Treatment) in the SOWS at the safety follow-up visit

10.2. Adverse Events

An adverse event (AE) is defined as any untoward medical occurrence in a patient or clinical investigation patient administered an IP that does not necessarily have a causal relationship with this treatment. An AE can therefore be any unfavorable and unintended sign (including an abnormal laboratory finding, for example), symptom, or disease temporally associated with the use of a medicinal product, whether related to the medicinal product, or not.

Pre-existing diseases or conditions will not be considered AEs unless there is an increase in the frequency or severity, or a change in the quality, of the disease or condition. Worsening of a pre-existing condition is considered an AE.

Events that occur in patients treated with placebo, or during treatment free periods of the study, are also considered AEs.

An AE is defined as treatment-emergent if the first onset or worsening is during or after the first administration of IP. AEs (serious and non-serious) that are ongoing at the last study visit should be followed for up to 28 days following the last dose or until one of the above conditions (resolution, etc.) is (are) met. In the case of missing or partial dates, an adverse event will be considered treatment-emergent unless the available data excludes that possibility.

An AE summary table will present totals for the following:

- All treatment-emergent adverse events (TEAEs)

[redacted]

- TEAEs by severity (mild, moderate, severe)
- TEAEs leading to study drug discontinuation
- TEAEs by relationship (not related, unlikely related, possibly related)
- SAEs

Summaries of incidence rates (frequencies and percentages) of individual AEs by MedDRA SOC and preferred term will be prepared. Such summaries will be displayed for all TEAEs, TEAEs by maximum severity, and TEAEs by strongest relationship to study drug.

Each patient will be counted only once within each summation level (SOC; preferred term). If a patient experiences more than one TEAE within each summation level, the TEAE with the strongest relationship or the maximum severity, as appropriate, will be included in the summaries of relationship and severity.

All AEs will be presented in a listing.

These summaries will include SAEs; additionally, they will be presented separately (see below).

10.3. Deaths, Serious Adverse Events, and Other Significant Adverse Events

Summaries of incidence rates (frequencies and percentages) of individual SAEs by MedDRA SOC and preferred term will be presented. Additionally, SAEs will be listed separately.

TEAEs leading to discontinuation of study drug will be analyzed and presented in a manner identical to the SAEs.

Adverse events potentially related to abuse identified via the Misuse, Abuse, Diversion Drug Event Reporting System (MADDERS®) will be presented in a separate listing.

All deaths will be reported in a listing.

10.4. Clinical Laboratory Evaluation

Laboratory data including Chemistry, Hematology, and urinalyses will be summarized by treatment and visit for the Safety Population. Descriptive summaries of actual values and changes from baseline will be presented by study visit for each clinical laboratory analyte and for each treatment group. 95% confidence intervals will be presented for change from baseline.

Laboratory values will be displayed in the data listings with their corresponding normal ranges, and those values that are outside the normal range will be flagged

Results of urine drug screens and pregnancy tests will be presented in listings.

10.5. Vital Signs, ECGs, Physical Findings, and Other Observations Related to Safety

10.5.1. Vital Signs

Descriptive summaries (mean, SD, median, minimum, and maximum) of actual values and changes from baseline at each assessment time point will be reported for vital signs including weight, body mass index, body temperature, pulse rate, systolic blood pressure and diastolic blood pressure. 95% confidence intervals will be presented for change from baseline. Vital sign values, including abnormal values, will be listed.

10.5.2. ECGs

Descriptive summaries (mean, SD, median, minimum, and maximum) of actual values and changes from baseline at each assessment time point will be reported for each ECG parameter. 95% confidence intervals will be presented for change from baseline.

The ECG interpretation will be performed by the central reader and confirmed by the Investigator, recorded in the CRF as normal, abnormal but not clinically significant, or abnormal and clinically significant. The number and percentage of patients in each category will be reported in shift tables vs the baseline interpretation at each visit.

ECG values, including interpretation, will be listed.

10.5.3. Physical Examinations

Physical examination data will be presented in a data listing

10.5.4. Other Safety Measures

10.5.4.1. Columbia Suicide Severity Rating Scale (C-SSRS)

Based on the C-SSRS results, the overall number of patients with lifetime and/or current suicidal ideation (by item and category), suicidal behavior (by item and category), or self-injurious behavior at the screening and baseline visit will be summarized by visit and treatment group. Additionally, the overall number of patients with any suicidal ideation or behavior (by type and in total) or self-injurious behavior while on-treatment will be provided by treatment group. Patients will only be counted once for on-treatment ideation and on-treatment behavior, at the worst-case response for each item.

10.5.4.2. Changes from baseline in the Sexual Functioning Questionnaire Short-Form (CSFQ-14)

The CSFQ-14 will be analyzed by gender. Note: there is a male version and female version for the CSFQ-14. Descriptive summaries (mean, SD, median, minimum, and maximum) of actual values and changes from baseline at each assessment time point and last on-treatment assessment will be calculated for each CSFQ domain score and the total. Note that the CSFQ is considered both an efficacy and safety endpoint.

10.5.4.3. Subjective Opiate Withdrawal Scale (SOWS)

The SOWS is a self-administered 16-item scale for evaluating opioid withdrawal symptoms. The SOWS will be administered at End of Treatment (Visit 5/ET) (after 6 weeks of treatment/ET), then completed at home at +1 day, +2 days, +3 days, +6 days, +9 days, and +12 after Visit 5/ET, and lastly administered at the Post Study Follow-up (Visit 6). Operationally, the timing of assessments are anchored around the last dose as entered into the EDC; in the case of entry errors of this field, assessments may be triggered on days differing from the above schedule. In these instances, the date of last dose will be used as anchor and available assessments will be windowed as follows:

Day	Window
End of Treatment	Last assessment on/ prior to last dose day, with priority given to the in-clinic assessment
Day +1	Day +1 from last dose
Day +2	Day +2 from last dose
Day +3	Day +3, 4 from last dose
Day +6	Day +5, 6, 7 from last dose
Day +9	Day +8, 9, 10 from last dose
Day +12	Day +11 from last dose and later
Follow-up Visit	In-clinic assessment at follow-up visit. This visit may also be utilized for one of the above visits if it falls in-window and there are no at-home assessments in-window.

If there is more than one assessment in-window with equal priority, then the one closest to the planned day will be used. In the case of ties, the set of questions with the highest total SOWS score will be used for tabulations (but all assessments will be listed).

Descriptive summaries (mean, SD, median, minimum, and maximum) of actual values and changes from baseline (Visit 5/ET) at each assessment time point will be reported for the total SOWS score and each individual question. 95% confidence intervals will be presented for change from baseline.

10.5.4.4. Misuse, Abuse, Diversion Drug Event Reporting System (MADDERS®)

The MADDERS® is a structured method to:

- Identify potential abuse-related events associated with study drug
- Classify the events according to standardized definitions

[redacted]

- Identify potential abuse-related events not already classified as an adverse event of interest or drug accountability discrepancy.

The Investigator at each site will be trained to identify related adverse events and drug accountability discrepancy. This information is collected throughout the study in real time. If any of these events are identified during the study, the Investigator will determine if it matches a MADDERS® triggering event. If so, the MADDERS® Supplemental Adverse Event Form or the MADDERS® Supplemental Drug Accountability Form will be completed. These events are adjudicated by a MADDERS® Adjudication Committee after the patient exits the study.

In addition to, all participants (regardless of the occurrence of a MADDERS® triggering events) will complete the MADDERS® Medication Use Survey on Visit 6. If the participant is lost to follow-up prior to Visit 6, the Medication Use Survey should be completed based on the site's knowledge of the participant.

The MADDERS® events will be reported in a separate appendix to the CSR and is outside the scope of this SAP.

The MADDERS® Medication Use Survey will be summarized by treatment with the count and percentage of patients with each response to each question.

11. OTHER ANALYSES

N/A

[redacted]

12. CHANGES TO THE ANALYSES PLANNED IN THE PROTOCOL

[redacted]

[redacted]

13. REFERENCES

Diggle P, Kenward M. Informative dropout in longitudinal data analysis (with discussion). Appl Stat. 1994;43:49-93

Lu K. and Mehrotra D.V. Specification of covariance structure in longitudinal data analysis for randomized clinical trials. Statistics in Medicine 2010 Feb 20; 29(4):474-88.

Rubin, D. B. (1976), “Inference and Missing Data,”Biometrika, 63, 581–592.

[redacted]

14. APPENDIX

14.1. Schedule of Events

Procedures and Evaluations	Screening	Baseline	Double-Blind Treatment Period			Post Study Follow-up
Visit	**1**	**2**	**3**	**4**	**5 or ET**	**6**
Study Week	**-5 to -1**	**0**	**2**	**4**	**6**	**8**
Study Day	**-35 to -7**	**1**	**15 ± 2**	**29 ± 2**	**43 + 2**	**57 + 7**
Screening						
Informed consent	X					
Optional Pharmacogenomics consent	X[g]					
Clinical Trial Subject Registry Consent	X[k]					
Placebo response training[a]	X	X	X	X	X	
Inclusion/exclusion criteria	X	X				
Psychiatric history (MINI 7.0.2)	X					
MSI-BPD	X					
Prior and concomitant medications	X	X	X	X	X	X
Demographics and medical history	X				X[l]	
Vital signs, weight, height[b], and BMI	X	X	X	X	X	X
Physical examination, including a brief neurological exam	X				X	
12-lead ECG	X	X			X	
Pregnancy test[c]	X	X	X	X	X	X
Urine drug screen[d]	X					
Clinical laboratory assessments	X			X	X	
PK				X	X	
Randomization		X				
Study Drug						
Dispense study drug		X[f]	X	X		
Study drug return, compliance, and accountability			X	X	X	
Optional Pharmacogenomics						
Blood sample collection		X[g]	*	*	*	*
Efficacy						
MADRS	X	X	X	X	X	
CGI-S		X	X	X	X	
CGI-I			X	X	X	
HARS		X	X	X	X	
SDS		X	X	X	X	
PGIC			X	X	X	
PGIS		X	X	X	X	
Safety						
AEs		X	X	X	X	X
MADDERS®			X[h]	X[h]	X[h]	X[i]
C-SSRS	X	X	X	X	X	X
CSFQ-14		X			X	
SOWS					X	X[j]

Abbreviations: AE = adverse event; ET = early termination; BMI = Body Mass Index; TSH=Thyroid Stimulating Hormone; CGI-I = Clinical Global Impression – Improvement scale; CGI-S = Clinical Global Impression – Severity scale; HARS = Hamilton Anxiety Rating Scale; CSFQ-14 = Changes in Sexual Functioning Questionnaire Short-Form; C-SSRS = Columbia-Suicide Severity Rating Scales; ECG = Electrocardiogram; MADRS = Montgomery-Åsberg Depression Rating Scale; MINI 7.0.2 = Mini International Neuropsychiatric Interview, Version 7.0.2; MSI-BPD = McLean Screening Instrument for Borderline Personality Disorder; PGIC = Patient Global Impression of Change; PGIS = Patient Global

[redacted]

Impression of Severity; PK = pharmacokinetic; SDS = Sheehan Disability Scale; MADDERS® = Misuse, Abuse, Diversion Drug Event Reporting System; SOWS = Subjective Opiate Withdrawal Scale.

a A placebo response reduction document is presented to patients by study staff for review and discussion at every in-clinic visit from Screening (Visit 1) through end-of Treatment (Visit 5/ET).
b Height to be measured at Screening (Visit 1) only.
c Serum pregnancy test at Screening (Visit 1) and on-site urine pregnancy test at all other visits for all women of childbearing potential, including women who have had bilateral tubal ligation. Serum pregnancy test may be conducted at Investigator's discretion at End of Treatment (Visit 5/ET).
d Performed at Screening (Visit 1) and only at other visits if Investigator suspects drug use or if a prescribed scheduled medication is washed out to participate in the study (eg., benzodiazepines, stimulants, opioids). At Screening, both an on-site urine dip-stick and a central laboratory sample will be obtained.
e Chemistry, Hematology, HbA1c and TSH at Screening. Chemistry and Hematology only at Visit 4 and Visit5.
f Patient will take first dose of study drug at the clinic at the end of Baseline (Visit 2).
g PGx sample ideally taken at Visit 2 but can be taken at any post-Baseline visit as well (*), noting that Baseline (Visit 2) is best due to potential for discontinuation before subsequent visits.
h MADDERS® questionnaires will be administered only if any pre-defined AEs and/or DADEs suspect for misuse, abuse or diversion event is identified, or any drug accountability issues above the pre-defined threshold are identified.
i MADDERS® Medication Use Survey will be administered to all patients at Visit 6.
j To be completed at +1 day, +2 days, +3 days, +6 days, +9 days, and +12 days at home after the Completion of Treatment, and then in-person at the Post Study Follow-up visit.
k The Clinical Trial Subject Registry (CTSR) is a mandatory part of the consent process. A site can elect not to participate in the CTSR, with approval from the sponsor, and there will be no protocol deviations recorded for leaving it out of the procedures.
l Collection of demographics at Visit 5/ET will only include re-collecting current nicotine, alcohol, and THC usage patterns post double-blind treatment.

14.2. Random Seeds

The following list of numbers will be used for random seeds where required for MI processes:

54748395

35852537

55264123

85945257

84116945

34285262

13458178

72845657

28932945

95785785

These will generally be used in order for the primary analysis, then secondary endpoints and sensitivity analyses. For cases where identical code may be applied to more than one outcome or a sensitivity analysis that uses a minor variation on the primary code, the second to last digit will be incremented by 1 to produce new seeds for the subsequent outcome/analysis. If a single dataset requires more than the 10 seeds above, additional seeds will be generated by incrementing the last digit by 1. All seeds used will be documented in the programming specifications and the programs themselves.

15. ATTACHMENTS

N/A

[redacted]