CLINICAL STUDY PROTOCOL: SP-624-201 **Study Title:**	A Multicenter, Double-Blind, Randomized, Placebo-Controlled Study of the Safety and Efficacy of SP-624 in the Treatment of Adults with Major Depressive Disorder
Sponsor:	Sirtsei Pharmaceuticals, Inc.
Phase of Development:	Phase 2
NCT Number	NCT04479852
Protocol Number:	SP-624-201
Statistical Analysis Plan:	Original
Statistical Analysis Plan Version Number:	1.0
Statistical Analysis Plan Version Date:	26 August 2022

[redacted]

Sirtsei Pharmaceuticals, Inc.

SP-624-201

A Multicenter, Double-Blind, Randomized, Placebo-Controlled Study of the Safety and Efficacy of SP-624 in the Treatment of Adults with Major Depressive Disorder

[redacted]

Document Revision History:

Version	Date	Content/Modifications
01	08/26/2022	Final Original

[redacted]

Statistical Analysis Plan

Approval

Signature Page

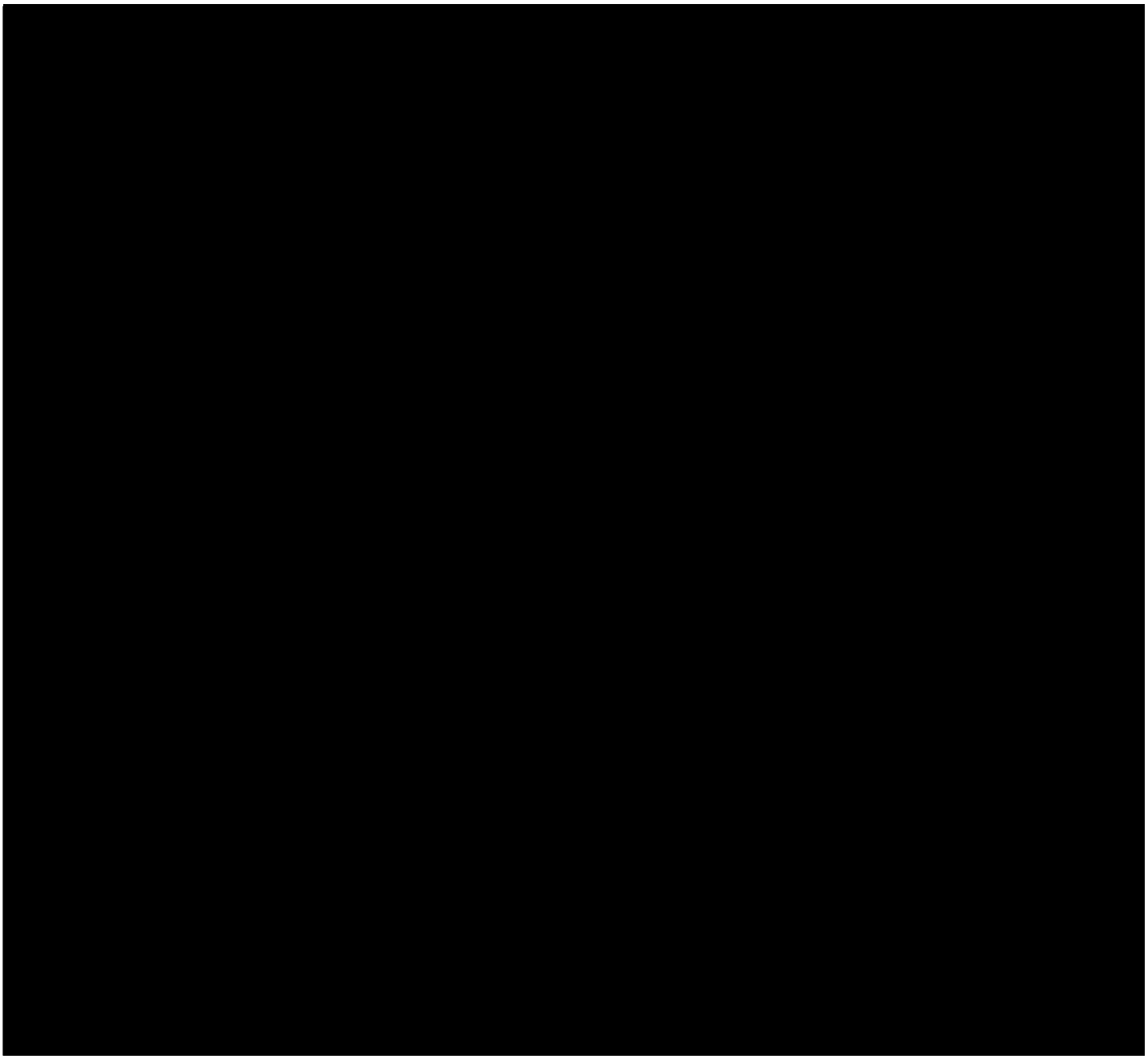

Table of Contents

1 Abbreviations

AE	Adverse Event
ANCOVA	Analysis of Covariance
ATC	Anatomical Therapeutic Chemical
BOCF	Baseline Observation Carried Foward
CI	Confidence Interval
CMH	Cochran-Mantel-Haenzel
ECG	Electrocardiogram
HAM-D	Hamilton Depression Rating Scale
ITT	Intent to Treat
LOCF	Last Observation Carried Forward
LTFU	Lost to Follow-up
MAR	Missing at Random
MADRS	Montgomery-Asberg Depression Rating Scale
MDD	Major Depressive Disorder
MedDRA	Medical Dictionary for Regulatory Activities
mITT	Modified Intent to Treat
MMRM	Mixed Model for Repeated Measures
NAD+	Nicotinamide-adenine dinucleotide coenzyme
PP	Per-Protocol Population
PT	Preferred Term
Q-LES-Q-SF	Quality of Life Enjoyment and Satisfaction Questionnaire
QIDS-SR	Quick Inventory of Depressive Symptomology
SAE(s)	Serious Adverse Event(s)
SAS®	Statistical Analysis Software®
SDS	Sheehan Disability Scale
SOC	System Organ Class
SD	Standard Deviation
[redacted]	
SOC	System Organ Class
SOE	Study Order of Events

2 Introduction

2.1 Background

SP-624 is a novel selective activator of [redacted] intended for the treatment of major depressive disorder, also known as unipolar or major depression. Major depressive disorder is characterized by a persistent feeling of sadness or a lack of interest in outside stimuli, with an estimated 16 million people living with major depression in the United States as of 2012. Antidepressant therapy aims to show symptom remission, often defined as at least 50% improvement from Baseline score on a depression scale such as the Hamilton Depression Rating Scale. However, there is room for improvement with many remitted patients reporting poorer quality of life and running a higher risk of depression relapse.

[redacted] Phase 1 clinical studies have shown that oral doses of SP-624 up to 30 mg/day were generally safe and well tolerated, without any reports of severe AEs, SAEs, deaths, or AEs leading to withdrawal from the study.

Sirtsei Pharmaceuticals, Inc. plans to investigate the safety and efficacy of SP-624 in this Phase 2, double-blind, placebo controlled, multicenter study in adults with moderate to severe MDD. The study will evaluate the safety and efficacy of a 4-week, multidose regimen of SP-624 versus placebo in approximately 300 subjects. Pending the efficacy and safety profile observed from this study, additional nonclinical and clinical studies will be conducted.

2.2 Changes to Planned Analyses

The protocol definition of the safety population is all subjects in the mITT population who receive at least one dose of study drug. This population definition has been corrected in section 7 below to all subjects that were randomized to treatment and received at least one dose of study drug. The updated safety population definition will be used for all safety analyses.

3 Study Objectives

The primary objective of this study is to evaluate the efficacy of SP-624 administered once-daily for up to 4 weeks compared to placebo in the treatment of adults with MDD. Additionally, there are two secondary objectives to evaluate the safety and tolerability of SP-624 administered once-daily for up to 4 weeks compared to placebo in the treatment of adults with MDD and to characterize the population PK of SP-624 in plasma when administered orally to subjects with MDD.

4 Investigational Plan and Study Design

This is a Phase 2, multicenter, double-blind, randomized, placebo-controlled study of the safety and efficacy of SP-624 in the treatment of adult subjects with MDD as defined by the DSM-5. Following the successful completion of a Screening Phase, subjects will be randomized to one of two treatment groups in a 1:1 ratio and will receive SP-624 or placebo over a treatment period of four weeks (Treatment Group 1: SP-624 20 mg/day; Treatment Group 2: Placebo).

Approximately 300 subjects in a 1:1 ratio (150 per treatment group) will be enrolled into the study. Each subject will complete Screening, Baseline, Treatment, and Follow-up Visits.

5 Sample Size Justification

Recent studies of MDD using change from Baseline in MADRS total scores suggest that placebo will yield an approximately nine unit decrease over four weeks with a standard deviation (SD) of around 9.2. SP-624 is expected to yield at least a 12 unit decrease in MADRS total score over four weeks with a

similar SD of around 9.2. A minimum of 150 subjects per treatment arm (300 total) would be needed to detect a treatment difference of three units with 80% statistical power. The sample size was estimated using NQuery 8 (V8.5.2.0) two group test of equal means (MTT0-1: www.statsols.com).

6 Visit Schedule and Visit Windows

A Screening period of up to 28 days (Day -28 to Day -2) and a Baseline Visit (Day -1) will precede initiation of the Treatment Period. During the Screening Period, all Screening assessments will be completed, and any current depression medications will be discontinued. All Screening assessments should be completed before discontinuing any current depression medications. Subjects who continue to satisfy all entry criteria at the Baseline Visit will be randomized in a 1:1 ratio to one of the two treatment groups. At the completion of the Baseline Visit, randomized subjects will be dispensed study drug and will be instructed to begin dosing the next morning (Day 1).

During the Treatment Period, subjects will be scheduled to return to the clinic for safety and efficacy assessments at the end of Weeks 1, 2, 3, and 4 as specified in the SOE (Table 2 of the protocol). After the final dose of study drug, subjects will complete a Follow-up Period of two weeks and will return at the end of Week 5 (1 week after final dose) and Week 6 (2 weeks after final dose). Subjects may be treated with anti-depressant medications according to physician recommendations after completion of the Week 5 Visit unless a clinically significant lab value is identified at the Week 4 Visit. If a clinically significant lab value is identified at the Week 4 Visit, consult with the Medical Monitor to determine if the lab value needs to be followed prior to initiating new therapy. Unless there are any outstanding safety issues that require follow-up, subjects will be discharged from the study at the Week 6 Visit.

Table 6.1, Sirtsei 201 Study Visit Schedule

Study Period	Screening/ Washout	Baseline	Treatment Period				Follow-up Period	
Study Visit	1	2	3	4	5	6	7	8
Study Day / Week	Day -28 to -2	Day -1	Week 1 (Day 7±3 Days)	Week 2 (Day 14±3 Days)	Week 3 (Day 21±3 Days)	Week 4 / Premature D/C (Day 28±3 Days)	Week 5/ 1-Week Follow-up (7 ±3 days post last dose)	Week 6/ 2-Week Follow-up (14 ± 3 days post last dose)
Event / Assessment								
Clinic Visit[1]	X	X	X	X	X	X	X	X
Informed Consent	X							
Inclusion/Exclusion	X	X						
Diagnosis Confirmation and Treatment History	X							
Height and Weight[2]	X	X				X		
Physical Exam[3]	X	X	X	X	X	X	X	
C-SSRS[4]	X	X	X	X	X	X	X	X
Pregnancy Test[5]	X	X				X		
Urine Drug Screen	X							
Clinical Safety Labs[6]	X	X		X		X	X	
Vital Signs[7]	X	X	X	X	X	X	X	X
12-Lead ECG[8]	X	X	X	X	X	X	X	
MADRS	X	X	X	X	X	X	X	
CGI-S	X	X	X	X	X	X	X	
HAM-D, SDS, QIDS-RS, Q-LES-Q-SF		X		X		X		
Randomization		X						
Dispensation / Collection / Accountability of Study Drug		X	X	X	X	X		
Plasma PK Blood Sampling		X	X	X		X		

Study Period	**Screening/ Washout**	**Baseline**	**Treatment Period**				**Follow-up Period**	
Study Visit	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Study Day / Week	**Day -28 to -2**	**Day -1**	**Week 1 (Day 7±3 Days)**	**Week 2 (Day 14±3 Days)**	**Week 3 (Day 21±3 Days)**	**Week 4 / Premature D/C (Day 28±3 Days)**	**Week 5/ 1-Week Follow-up (7 ±3 days post last dose)**	**Week 6/ 2-Week Follow-up (14 ± 3 days post last dose)**
[redacted]								
Prior and Con Med Record	X	X	X	X	X	X	X	X
AE Monitoring	X	X	X	X	X	X	X	X

Abbreviations: AE=adverse event; con med=concomitant medication; CGI-S=clinical global impression-severity; C-SSRS=Columbia-Suicide Severity Rating Scale; ECG=electrocardiogram; MADRS=Montgomery-Asberg Depression Rating Scale; PK=pharmacokinetics; Premature D/C=Premature Discontinuation from treatment; QIDS-SR=Quick Inventory of Depressive Symptomology – Self Report; Q-LES-Q-SF=Quality of Life Enjoyment and Satisfaction Questionnaire – Short Form; SDS=Sheehan Disability Scale

[1] If a clinic visit is not possible (e.g., due to illness), a remote visit may be conducted to complete all assessments that can be done remotely and assess for adverse events and concomitant medications. The Screening, Baseline, and Week 4/Premature Discontinuation Visits cannot be done remotely. The sponsor should be contacted prior to conducting any remote visits whenever possible.

[2] Height and body weight at Visit 1. Body weight only at Visit 2 and Visit 6/Premature Discontinuation.

[3] Physical examinations will include a review of the following body systems: general appearance, skin and extremities, ear/eyes/nose/throat, head/neck, heart/chest, abdomen, musculoskeletal, and neurological.

[4] Baseline/Screening Version at Visit 1; Since Last Visit Version at Visits 2-8.

[5] All females; Serum pregnancy test at Visit 1 and Visit 6 (or Premature Discontinuation); Urine pregnancy test at Visit 2 – if urine pregnancy test at Visit 2 is positive, a confirmatory serum pregnancy test will be conducted.

[6] Clinical safety labs include clinical chemistry, hematology, coagulation, and urinalysis. Labs should be collected in a fasting state. Urinalysis not required at Visit 4 (Week 2 visit).

[7] Vital signs include blood pressure, pulse rate, and oral body temperature. Vital signs to be assessed in a sitting position after three mins or more of quiet sitting.

[8] Two ECGs to be collected at the Baseline Visit separated by approximately 10 minutes. One ECG collected at all other visits.

7 Study Populations

The following populations will be considered for this study:

- All Subjects Population – all subjects who sign informed consent and are assigned a unique subject identification number. This population will be used to summarize disposition.

- Modified Intent-to-Treat (mITT) Population – all subjects who are randomized to receive treatment, receive at least one dose of study drug, and have a Baseline and at least one post-dose efficacy assessment. This population will be the primary population used to summarize demographics, baseline characteristics, medical history, treatment history, and all efficacy outcomes. Subjects will be analyzed according to randomized treatment group.

- Safety Population — all subjects randomized to treatment and who receive at least one dose of study drug. Safety subjects will be summarized according to the actual treatment they received. This population will be used to summarize all safety assessments.

- PK Population— all subjects who receive active study drug and have evaluable PK data based on actual treatment received, protocol compliance, adequate numbers of samples, and successful sample assays.

8 Statistical Methods

8.1 General Reporting Conventions

Study data will be provided in listings and sorted by subject number, study period, and assessment time. The study results will be displayed using descriptive summary tables and inferential analyses completed as described below. Summaries will be displayed by group for all tables and, optionally, for all subjects for baseline tables. For measurements that have multiple intermediate visits, results for observed data will be summarized for each time.

Standard numeric descriptive statistics include number of subjects or records observed (N), mean, standard deviation (std), median, minimum (min), and maximum (max) values. Standard categorical descriptive statistics include the count and percentages of subjects with a level of the variable summarized. Hypothesis tests will assess the null hypothesis that there is no difference in the treatment arms. Two-sided p-values of less than or equal to 0.05 will be considered statistically significant.

All summaries, statistical analyses, and individual subject data listings described below will be completed using Version 9.4 or later of the SAS Statistical Analysis System (SAS Institute, Inc. Cary, NC).

Subjects who are randomized but not treated and those who are treated but have no post-treatment follow-up assessments or data will be listed, but not analyzed. For all listings with dates, the relative study day will be presented, and a minus sign will denote study day is prior to first dose (day 1) while a plus sign will denote number of study days after last dose. Days occurring during treatment period will use neither a plus nor a minus sign.

8.1.1 Efficacy Assessment Windowing

For all efficacy analyses, windowing will be applied to ensure study day of assessment does not fall too far from scheduled assessment day. Assessments will be considered part of the scheduled visit window if they fall within +/- 3 days of the nominal study visit day. In the event of multiple assessments within the same visit window, the assessment closest to the nominal study day will be used. In the event of a tie, the

earliest study day will be used. A visit must be completed within 2 days of last treatment dose to be included as a treatment period visit. Visits completed 3 or more days after last dose will be excluded from the treatment period analysis. Follow-up period windowing will be relative to date of last dose of study drug: assessments completed 3 to 10 days after the last dose of study drug will be windowed to the Week 5/1-Week Follow-up Visit.

Efficacy assessment listings will include both the reported and windowed visit labels. Any assessment completed outside of the windows for which that particular assessment was scheduled to be completed per the protocol will be labeled as "Out of Window".

8.1.2 Safety Assessment Windowing

For all safety summaries, windowing will be applied to the Week4/Premature Discontinuation Visit and Follow-up Period Visit assessments to ensure safety data from subjects with similar lengths of exposure and/or follow-up are summarized together. Windowing of safety assessment from these visits will be applied as follows:

- Safety assessments will be considered part of the scheduled visit window if they fall within +/- 3 days of the nominal study visit day.
- Safety assessments completed on Days 1-3 will also be windowed to the Week 1 visit.
- Safety assessments completed 3 to 10 days after the last dose of study drug will be windowed to the Week 5/1-Week Follow-up Visit
- Safety assessments completed 11 to 17 days after the last dose of study drug will be windowed to the Week 6/2-Week Follow-up Visit
- In the event of multiple assessments within the same visit window, the assessment closest to the nominal study day will be used. In the event of a tie, the latest study day will be used.

Subjects missing safety assessments at baseline will have the most recent visit value prior to date of first dose used as the baseline value.

Safety assessment listings will include both the reported and windowed visit labels. Any assessment completed outside of the windows for which that particular assessment was scheduled to be completed per the protocol will be labeled as "Unscheduled".

8.1.3 Relative Day

For listings that include visit/event dates, the relative day will also be listed. The relative day will be expressed using the following methodology:

- Date of first dose of study drug is Relative Day 1.
- No Day 0 will be used.
- For visit/event dates that occurred before the date of the first dose of study drug, the Relative Day will be expressed with a minus (-) sign followed by the number of days before the date of first dose of study drug that correspond to the date of the visit/event.
- For visit/event dates that occurred after the date of the last dose of study drug, the Relative Day will be expressed with a plus (+) sign followed by the number of days after the date of the last dose of study drug that correspond to the date of the visit/event.
- Relative Days shown without a minus or plus sign will correspond to the day to dosing.

Assessment windowing, as described in Sections 8.1.1 and 8.1.2, will be performed using the Relative Day.

8.2 Adjustments for Interim Analyses and Multiplicity of Endpoints

The study has a single primary effectiveness endpoint and there are no planned interim analyses; hence, no adjustment for multiplicity will be implemented.

8.3 Evaluations and Statistical Analyses

8.3.1 Subject Disposition

Subject disposition will include total number of subjects summarized by number screened and number of screen failures if they are recorded in the database. Summaries will include the number of subjects randomized by treatment group and overall. Summaries for randomized subjects who completed the study treatment period and as well as those who withdrew from the study treatment period will be summarized by number of subjects and percentages of subjects. Summaries for randomized subject who completed the follow up period and as well as those who withdrew from the follow up period will be summarized by number of subjects and percentages of subjects. Finally, the reasons for withdrawing from the study treatment period and follow up period will be summarized. Totals and percentages of subjects in each analysis dataset will be tabulated by treatment arm.

8.3.2 Protocol Deviations

Protocol deviations will be divided into minor protocol deviations and major protocol deviations, per Investigator and summarized using the safety population. For each protocol deviation type and study event, the count and percentage of subjects with that protocol deviation will be summarized by treatment group and overall. The total number of protocol deviations of that type will also be presented by treatment group and overall.

Protocol deviations will be listed using the all subjects population.

8.3.3 Demographics and Baseline Characteristics

Demographics information will be summarized using descriptive statistics. For gender, race, and ethnicity the summarization of those categorical values will be performed using percentages and frequency. Age (at Screen), height, weight (at Screen and Baseline), and BMI (at Screen and Baseline) will be summarized using descriptive statistics as total number of subjects surveyed (N), mean, median, minimum, and maximum.

Listings and summaries of demographics, body measurements at Screen and Baseline, medical history, depression and current depressive episode history, and current depressive episode treatment history will be provided for the mITT or safety populations.

8.3.3.1 Concomitant Medications

Concomitant medications will be tabulated by major therapeutic class and coded medication name, but no formal analyses will be performed. Number and percentage of subjects will be presented by treatment group and overall. Concomitant medication summaries will be presented separately for the following:

- Concomitant Medications taken during the Treatment Period (ie, either taken or continuing from date of first dose of study drug to date of last dose of study drug for each subject); and
- Concomitant Medications taken during the Follow-up Period (ie, either taken or continuing after date of last dose of study drug for each subject).

Medications taken and stopped prior to first dose of study drug (ie, prior medications) will be included in the listing, but will not be summarized as a concomitant medication.

8.3.4 Treatment Administration and Compliance

Descriptive numerical summaries (N, mean, standard deviation, median, minimum, and maximum) will be presented for both the total drug exposure in days and total drug exposure in number of capsules per day by treatment group and overall. The percent compliant with prescribed dosing during the treatment period will similarly be summarized descriptively. Counts and percentages of subjects that met total days of study drug exposure marks as well as overall total days of study drug exposure will be summarized

categorically by study week. Counts and percentages will also be shown for number of subjects by total percent compliance with prescribed dosing during the treatment period and number of subjects for whom the prescribed dose was reduced from two capsules a day to one capsule a day.

8.3.5 Primary Estimand

The primary efficacy estimand, is comprised of the primary endpoint, target population, intercurrent events of interest and strategy for addressing them, analysis approach to compare treatments and summary measure (hypothesis to be tested).

8.3.5.1 Primary Efficacy Endpoint

The primary efficacy endpoint for this study is the change from Baseline to Week 4 in total MADRS score. Efficacy analyses will be conducted using the mITT population, assuming subjects had dosed as directed for 4 weeks.

8.3.5.2 Intercurrent Events of Interest and Strategy for Addressing

The intercurrent events of interest and strategy for addressing is given in the following table:

Table 8.1, Handling of Intercurrent Events

Estimand	Intercurrent Event	Rule / Rationale
Primary	Premature discontinuation / Permanent treatment discontinuation	Hypothetical "as own treatment" strategy. Clinical course is the same as other subjects in the same treatment group, who had not discontinued study treatment. Exclude observations that occur more than 3 days after the last dose of study drug for subjects who discontinue. The objective of the study is to evaluate the effect of SP-624 subjects who are provided drug to take on a standard schedule up to 4 weeks. After 3 days, the drug should be cleared from all subjects, so the inclusion of any such data point is inconsistent with the goal of the estimand.
	Drug non-compliance / Missed study medication	Treatment policy strategy. Some degree of non-compliance missed doses would be expected to occur in standard use.
	Protocol deviations not leading to study discontinuation	Treatment policy strategy Overall, if not sufficient to lead to subject removal, the expectation is that they would not significantly impact subject results.
	Concomitant medications in violation of the protocol	Treatment policy strategy This is expected to be a rare event in this clinical trial and treatment effect will be estimated irrespective of concomitant medication violations.

The hypothetical "as own treatment" strategy was selected for premature discontinuation or treatment discontinuation intercurrent events to evaluate the effect of treatment with respect to its pharmacologic efficacy. For the change from week 4 to week 5 follow-up secondary analysis, only subjects that were included in the week 4 primary analysis will be used. This should provide the most unbiased data for this

phase 2 study to inform future studies. The treatment policy strategy will be used for all other defined intercurrent events.

8.3.5.3 Primary Efficacy Hypothesis

The null hypothesis is that the mean change from Baseline to Week 4 in total MADRS score will be the same in both groups. The following hypothesis statement to be tested for the comparison of the primary efficacy endpoint is:

$$H_0: \mu_{SP\text{-}624} = \mu_{Control} \text{ versus } H_1: \mu_{SP\text{-}624} > \mu_{Control}$$

Where $\mu_{SP\text{-}624}$ is the mean change in MADRS score from baseline to week 4 for the SP-624 group and $\mu_{Control}$ is the mean change for the control group.

8.3.5.4 Primary Efficacy Analysis

The primary efficacy hypothesis will be evaluated by comparing SP-624 to placebo using the mean change from baseline in total MADRS total score after 4 weeks of treatment using a Mixed Model for Repeated Measures (MMRM). The analysis model will include fixed effects for treatment, week and the interaction of week by treatment, and a random effect of subject within treatment. Baseline MADRS total score will be included as a covariate in the model. The treatment difference in terms of the mean change from baseline to a given time point will be estimated and tested from this model. An unstructured covariance matrix will be used to model the correlation among repeated measures. The Kenward-Rogers approximation will be used for the denominator degrees of freedom. The criterion for declaring SP-624 superior to placebo for MADRS total score change from baseline is that the p-value is less than or equal to 0.05.

In the event of a model which fails to converge, the following covariance structures will be considered (in order) until convergence is obtained:

1. Heterogenous Toeplitz
2. Toeplitz
3. First order auto-regressive [AR(1)]
4. Compound symmetry

If a covariance structure other than Type=UN is used, a sandwich estimator using the EMPIRICAL option on the PROC MIXED statement will be used.

8.3.5.5 Sensitivity and Supporting Analyses

A dropout rate of 20-30% is typically seen in MDD studies. Subjects will be grouped by the visit at which they had their last MADRS total score measured. This will result in five categories of discontinuation: Week 1 dropouts, Week 2 dropouts, Week 3 dropouts, Week 4 dropouts, and completers. Mean change from Baseline in MADRS total score will be plotted by the dropout category and by reason for discontinuation, to assess whether this efficacy measure appears to be correlated with study dropout.

The mechanisms that cause missing data may be missing at random (MAR) or may be missing not at random (MNAR). The MMRM model (as implemented within the PROC MIXED procedure in SAS) used in the primary analysis of the primary efficacy endpoint relies on the assumption of MAR.

Therefore, a sensitivity analysis using the tipping-point approach may be used to assess the robustness of the primary analysis approach, in the event the primary analysis is positive. Missing data are first imputed for all visits under the missing-at-random (MAR) assumption, and then a worsening/shift is applied. This is repeated until the result is no longer statistically significant (i.e., $p>0.05$). The tipping point (smallest worsening/shift in which the significant result turns non-significant) provides a measure of robustness of the primary result. Additionally, in the event the primary analysis is positive a multiple

imputation analysis will be performed with 100 imputations of the study data using available results and variables used in the primary analysis (week, treatment, week to treatment interaction, and baseline MADRS score) to model missing data for each treatment group. Each imputed dataset will be analyzed separately using the same MMRM model in the primary efficacy analysis. The estimates from the imputed model analyses will be combined for statistical inference.

8.3.5.6 Covariate and Other Supporting Analyses

The mean change from baseline in total MADRS total score after 4 weeks of treatment between treatment groups will be compared using the same MMRM model as the primary efficacy analysis but with age added to the model as a covariate analysis.

A primary concern for this study is the homogeneity of treatment groups across sites. A test for homogeneity using the same MMRM model as the primary efficacy analysis will be used with site and the interaction of site by treatment added as fixed effects. Sites with less than 10 subjects will be pooled to obtain at least 10 subjects. Sites will be pooled in order of site number. The last un-pooled site will be pooled with the last set of sites pooled if the site has less than 10 subjects.

8.3.6 Secondary Efficacy Endpoints

Seven secondary efficacy endpoints will be analyzed using the mITT populations. Where a model similar to the primary analysis is identified, the model may be reduced, if necessary, for convergence issues. These will be supported by descriptive summaries of results by group. These include:

- Change from Baseline to Weeks 1, 2, and 3 in MADRS total score for SP-624 vs. placebo. The outcome will be analyzed using the same methods as the primary outcome using estimate statement statements for each timepoint from the overall model.
- CGI-S is measured on a 7-point scale where 1 represents “normal” and 7 represents “most extremely ill”. A summary of percentage of subjects with each score at Baseline and at Weeks 1, 2, 3, and 4 will be presented. The change from Baseline in CGI-S can take on interval values between -6 (improvement from 7 to 1) and 6 (worsening from 1 to 7). Treatment comparison of change from Baseline to Weeks 1, 2, 3, and 4 in CGI-S scores will be performed using an MMRM model similar to that of the primary efficacy variable.
- Change from Baseline to Week 5/1-Week Follow-up and change from Week 4 to Week 5 will be assessed for the MADRS and CGI-S. Treatment comparisons will be performed using an MMRM model similar to that of the primary efficacy variable.
- The HAM-D-17 is a 17-item scale whose total score can range from 0 to 52. Change from Baseline to Week 2 and Week 4 in the 17-item Hamilton Depression Rating Scale (HAM-D-17) total score will be assessed for treatment differences using an MMRM model similar to that of the primary efficacy variable.
- The Sheehan Disability Scale (SDS) is a 3-part scale that measures the degree of disruption on work, social and family life using an 11-point scale where 0 represents “no disruption” and 10 represents “extreme disruption”. In addition to the 11-point scale, subjects are asked to indicate the number of days in the past week that were “lost” and numbers of days that were “unproductive”. The results of these questions have a range from 0 to 7. A total global functioning impairment score can be utilized by summing the scores from work, social and family life scales for a value range from 0 to 30. Treatment difference in change from Baseline to Week 2 and Week 4 for both global functioning impairment and days lost/unproductive will be assessed using an MMRM model similar to that of the primary efficacy variable. Change in days lost will be analyzed separately from change in unproductive days.
- The Quick Inventory of Depressive Symptomology – Self Report (QIDS-SR), is a 16-item self-reported scale where each item has a 4-point scale where 0 represents least impact scores while 3 represents greatest impact scores. Scale 6 and 7 are linked such that only one of these will be

scored. Likewise, scales 8 and 9 are linked such that only one of them will be scored. Nine domains are scored based on the highest score of items within that domain. The total score from all domains can take a value ranging from 0 to 27. Treatment difference in change from Baseline to Week 2 and Week 4 in the QIDS-SR will be assessed using an MMRM model similar to that of the primary efficacy variable.

- The Quality of Life Enjoyment and Satisfaction Questionnaire – Short Form (Q-LES-Q-SF) is a 16-item satisfaction scale where each item has a 5-point scale. A score of 1 represents "very poor satisfaction", while a score of 5 represents "very good satisfaction". Only the first 14 items are summed for a total score that ranges from 14 to 70 and is expressed as a percentage based on the maximum total score of the items completed (0– 100). Treatment difference in the change from Baseline to Week 2 and Week 4 in total scores will be assessed using an MMRM model similar to that of the primary efficacy variable. The last two items pertaining to satisfaction with medications and overall contentment will be summarized separately by treatment group.

8.3.6.1 Handling of Missing Items Within a Clinical Assessment

The final scores of the Q-LES-Q-SF, QIDS-SR and SDS are all constructed from multiple sub-questions within each assessment. It is possible that one or more sub-questions may be missing within each assessment. In this event, the last recorded score for this sub-question(s) may be carried forward from the most recent post-dose visit (Last Observation Carried Forward approach). In a similar fashion, missing baseline values may be carried over from the most recent screening visit, if available. Baseline/screening values will never be carried forward to impute missing post-dose values. An individual sub-question will not be carried forward for more than one visit. If the same sub-question is missing two visits in a row, then LOCF will be applied to the first missing visit and the sub-question will remain as missing for the second visit (with the total score to then be computed as missing). Further, the total score will be computed as missing if too many sub-questions, prior to applying the LOCF approach, are missing (>50%). This single imputation approach allows the total score to be calculated using the strength of the other sub-questions collected at that time, for that subject. The Sponsor believes this approach to be more accurate than either setting the entire score to missing or to imputing the worst possible score.

8.3.7 Safety Evaluations

The following are all safety evaluations for this study:

- Adverse Events (AEs)
- Clinical safety laboratory tests (serum chemistry [including thyroid function], hematology, coagulation, and urinalysis tests)
- Vital signs (blood pressure, pulse rate, and oral body temperature)
- Body weight
- 12-lead ECGs
- Physical examinations
- C-SSRS

8.3.7.1 Safety Analyses

All summaries of safety and tolerability will be performed on the safety population.

Laboratory parameters, vital sign data, and ECG results will be summarized by treatment group and presented in tabular and graphic formats where appropriate. Appropriate change from Baseline calculations for the mean change from baseline within each treatment group where appropriate will be used for laboratory parameters, vital sign values, ECG results, and C-SSRS results, but no formal statistical tests are planned.

Shift tables will also be provided for laboratory parameters.

Unscheduled and Repeat safety assessments will be listed but will not be summarized in the tabular summaries. Local laboratory results will be listed but will not be summarized in the tabular summaries.

8.3.7.2 Adverse Events

An AE is any untoward medical occurrence in a clinical study subject, whether or not the event is considered causally related to the study drug. An AE can be a new occurrence or an existing process that increases significantly in intensity or frequency.

All Aes must be recorded in the appropriate section of the CRF.

Adverse event monitoring for each subject begins upon Screening and continues until study completion. Abnormalities present at Screening are considered AEs only if they reoccur after resolution or worsen in intensity or frequency during the study. The events recorded with a start date and time on or after the date and time of the first dose of study drug will be considered Treatment Emergent Aes (TEAEs). Each AE is to be classified as serious or non-serious, and categorized with respect to relationship to study drug, and intensity.

Summaries of the number and percentage of subjects in the safety population with at least one treatment-emergent AE (TEAE), classified according to preferred term and/or body system using the MedDRA dictionary, will be provided overall and by treatment group for:

- Overall summary
- All TEAEs
- TEAEs by relationship (not related, possibly related, probably related)
- TEAEs by severity (mild, moderate, severe)
- SAEs
- TEAEs leading to withdrawal from the study
- TEAEs leading to study drug dose reduction
- TEAEs leading to study drug dose taken with food
- Analyses of Aes with time of onset beginning 24 hours after the final dose through up to 28 days after the final dose.

In addition, a summary will be provided of the number and percentage of subjects in the safety population with at least one pre-treatment AE (i.e., Aes with date and time of onset prior to the first dose of study drug).

Listings of events leading to discontinuation or death will be provided.

No formal statistical tests are planned for AE data.

8.3.7.3 Vital Signs and ECGs

Vital signs and ECG results will be summarized by treatment group and visit and presented in tabular and graphic formats where appropriate. Descriptive summaries of the change from Baseline calculations will be provided for the mean change from baseline to detect potentially significant changes in vital sign values or ECG results. No formal statistical tests are planned between groups.

Tabular summaries of the following ECG results will also be provided:

- Number and percentage of subjects in each treatment group who meet criteria for maximum QTcF Interval (msec) and change from Baseline results;
    - A listing of ECGs for subjects in each treatment group who met at least one QTcF Interval criterion will also be provided;

- Number and percentage of subjects in each treatment group who meet criteria for maximum and minimum ECG-determined Heart Rate and maximum change from Baseline results; and
    - A listing of ECGs for subjects in each treatment group who met at least one ECG-determined Heart Rate criterion will also be provided.

8.3.7.4 Clinical Laboratory Assessments

Clinical laboratory test results for chemistry (including thyroid function), hematology, coagulation, and urinalysis will be summarized by treatment group and presented in tabular and graphic formats where appropriate. Descriptive summaries of the change from baseline calculations will be provided for the mean change from baseline within each group will be provided, but no formal statistical tests are planned between treatment groups. Shift tables will also be provided for laboratory parameters. Any clinically significant abnormal findings or changes will be captured as Aes.

Pregnancy test and urine drug screen test results will be listed for the safety population but not summarized.

8.3.7.5 Physical Exams

Physical examination findings will be listed for each subject at each assessment period. Clinically significant new findings and adverse changes in physical examination findings will be captured as Aes, therefore eliminating the need to summarize over time.

8.3.7.6 C-SSRS

Two summary tables will be prepared by treatment group for the C-SSRS assessment

1. Summary of C-SSRS (Baseline/Screening Version) at Screen
2. Summary of C-SSRS (Since the Last Visit Version) by Visit

8.3.8 PK Analyses

Plasma concentration data for SP-624 will be listed.

If data allow, results will be used for other analyses (e.g., PopPK) which will be presented in a separate report.

The population PK analysis plan will be prepared separately.

Plasma concentration data will remain blinded until the unblinding of the clinical database at the end of the study.

[redacted]

9 References

None.

10 List of Tables, Figures and Listings

10.1 List of Tables

Table Number	Title	Population	Included in Top Line Results Package
14.1.1	Subject Accountability	All Subjects	Yes
14.1.2.1	Subject Disposition	All Subjects	Yes
14.1.2.2	Protocol Deviations	Safety	
14.1.3	Demographics	Safety	Yes
14.1.4	Body Measurements at Screen and Baseline	Safety	
14.1.5	Medical History	Safety	
14.1.6.1	Concomitant Medications during the Treatment Period	Safety	
14.1.6.2	Concomitant Medications during the Follow-up Period	Safety	
14.1.7.1	Depression and Current Depressive Episode History	mITT	
14.1.7.2	Current Depressive Episode – Current Treatment at Screen	mITT	
14.1.7.3	Current Depressive Episode – Previous Treatments	mITT	
14.1.8	Treatment Administration and Compliance	Safety	Yes
14.2.1.1	Primary Efficacy Analysis Change from Baseline MADRS Score at Week 4 Using Mixed Models for Repeated Measures	mITT	Yes
14.2.1.2	Analysis of the Change from Baseline Total MADRS Score at Week 4 Type 3 Model Effects	mITT	
14.2.2.1	Tipping Point Analysis of Primary Efficacy Endpoint	mITT	
14.2.2.2	Multiple Imputation Analysis of Primary Efficacy Endpoint	mITT	
14.2.2.3	Covariate Analysis Effect of Age on Primary Efficacy Endpoint	mITT	
14.2.2.4	Test for Homogeneity of Site on Primary Efficacy Endpoint	mITT	
14.2.3.1	Change from Baseline to pre-Week 4 Timepoints in MADRS Score Using MMRM	mITT	Yes
14.2.3.2	Analysis of the Change from Baseline Total MADRS Score pre-Week 4 Timepoints Type 3 Model Effects	mITT	
14.2.4.1	Change from Baseline to Week 5 in MADRS Score and Change	mITT	Yes

Table Number	Title	Population	Included in Top Line Results Package
	from Week 4 to Week 5 in MADRS Score		
14.2.4.2	Analysis of the Change from Baseline to Week 5 and Week 4 to Week 5 Total MADRS Score Type 3 Model Effects	mITT	
14.2.5.1	Change from Baseline CGI-S Score Over Time	mITT	Yes
14.2.5.2	Analysis of the Change from Baseline CGI-S Score Over Time Type 3 Model Effects	mITT	
14.2.6.1	Change from Baseline to Week 5 in CGI-S Score and Change from Week 4 to Week 5 in CGI-S Score	mITT	Yes
14.2.6.2	Analysis of the Change from Baseline to Week 5 and Week 4 to Week 5 CGI-S Type 3 Model Effects	mITT	
14.2.7.1	Change in HAM-D-17 Score Over Time	mITT	Yes
14.2.7.2	Analysis of the Change from Baseline HAM-D-17 Score Over Time Type 3 Model Effects	mITT	
14.2.8.1	Change in SDS Score Over Time	mITT	
14.2.8.2	Analysis of the Change from Baseline SDS Score Over Time Type 3 Model Effects	mITT	
14.2.9.1	Change in QIDS Score Over Time	mITT	
14.2.9.2	Analysis of the Change from Baseline QIDS Score Over Time Type 3 Model Effects	mITT	
14.2.10.1	Change in Q-LES-Q-SF Score Over Time	mITT	
14.2.10.2	Analysis of the Change from Baseline Q-LES-Q-SF Score Over Time Type 3 Model Effects	mITT	
14.3.1.1	Overall Summary of Adverse Events	Safety	Yes
14.3.1.2.1	Treatment Emergent Adverse Events by SOC and PT	Safety	Yes
14.3.1.2.2	Treatment Emergent Adverse Events by Preferred Term	Safety	Yes

Table Number	Title	Population	Included in Top Line Results Package
14.3.1.3	Treatment Emergent Adverse Events by Relationship to Study Drug	Safety	
14.3.1.4	Treatment Emergent Adverse Events by Severity	Safety	
14.3.1.5	Treatment Emergent Serious Adverse Events by PT	Safety	
14.3.1.6	Treatment-Related Treatment Emergent Serious Adverse Events by PT	Safety	
14.3.1.7	Treatment Emergent Adverse Events Leading to Study Withdrawal	Safety	
14.3.1.8	Treatment Emergent Adverse Events Leading to Study Drug Dose Reduction	Safety	
14.3.1.9	Treatment Emergent Adverse Events Leading to Study Drug Taken With Food	Safety	
14.3.1.10	Adverse Events During Follow-up (with Onset 1 to 28 Days After Final Dose)	Safety	
14.3.2.1.1	Vital Signs and Weight	Safety	
14.3.2.2.1	Summary of ECG Results	Safety	
14.3.2.2.2	Maximum OTcF Interval (msec) and Change from Baseline Categorical Summary	Safety	
14.3.2.2.3	Maximum and Minimum ECG Determined Heart Rate (bpm) and Maximum Change from Baseline Categorical Summary	Safety	
14.3.2.3.1	Summary of Serum Chemistry	Safety	
14.3.2.3.2	Shift table of Serum Chemistry	Safety	
14.3.2.3.3	Summary of Thyroid Function Test Results	Safety	
14.3.2.3.4	Thyroid Function Tests Results Shift Over Time	Safety	
14.3.2.4.1	Summary of Hematology	Safety	
14.3.2.4.2	Shift table of Hematology	Safety	
14.3.2.5.1	Summary of Coagulation	Safety	
14.3.2.5.2	Shift table of Coagulation	Safety	
14.3.2.6.1	Summary of Urinalysis	Safety	
14.3.2.6.2	Shift table of Urinalysis	Safety	
14.3.2.7.1	Summary of C-SSRS (Baseline/Screening Version) at Screen	Safety	

Table Number	Title	Population	Included in Top Line Results Package
14.3.2.7.2	Summary of C-SSRS (Since the Last Visit Version) by Visit	Safety	

10.2 List of Figures

Figure Number	Title	Population	Included in Top Line Results Package
14.2.1.1	Change from Baseline in LSM MADRS Total Score Over Time by Treatment Group - Using MMRM	mITT	Yes
14.2.1.2	Change from Baseline in Observed Mean MADRS Total Score Over Time by Treatment Group	mITT	Yes
14.2.2.1	Change from Baseline in LSM HAM-D-17 Total Score Over Time by Treatment Group – Using MMRM	mITT	Yes
14.2.2.2	Change from Baseline in Observed Mean HAM-D-17 Total Score Over Time by Treatment Group	mITT	Yes
14.2.3.1	Change from Baseline in LSM CGI-S Score Over Time by Treatment Group – Using MMRM	mITT	Yes
14.2.3.2	Change from Baseline in Observed mean CGI-S Score Over Time by Treatment Group	mITT	Yes

10.3 List of Listings

Listing Number	Title	Population	Included in Top Line Results Package
16.1.1.1	Subject Disposition	All Subjects	Yes
16.1.1.2	Subject Eligibility and Randomization Information	All Subjects	Yes
16.1.2	Protocol Deviations	All Subjects	
16.1.3	Demographics	mITT and Safety	Yes
16.1.4	Body Measurements at Screen and Baseline	mITT and Safety	
16.1.5	Medical History	mITT and Safety	

Listing Number	Title	Population	Included in Top Line Results Package
16.1.6	Prior and Concomitant Medications	mITT and Safety	
16.1.7.1	Depression and Current Depressive Episode History	mITT and Safety	
16.1.7.2	Current Depressive Episode Treatment – Current at Screen and Previous	mITT and Safety	
16.1.8	Study Drug Prescription and Dosing Record	mITT and Safety	Yes
16.1.9	Investigator Comments	mITT and Safety	
16.2.1	MADRS Total Score Over Time	mITT	Yes
16.2.2.1	Individual Item MADRS Scores Over Time SP-624 Subjects (Part 1)	mITT	Yes
16.2.2.2	Individual Item MADRS Scores Over Time Placebo Subjects (Part 2)	mITT	Yes
16.2.3	CGI-S Scores Over Time	mITT	Yes
16.2.4	HAM-D-17 Scores Over Time	mITT	Yes
16.2.5	SDS Scores Over Time	mITT	
16.2.6	QIDS Scores Over Time	mITT	
16.2.7	Q-LES-Q-SF Scores Over Time	mITT	
16.3.1	Adverse Events	Safety	Yes
16.3.2.1	Vital Signs by Visit	Safety	
16.3.2.2	Body Measurements by Visit	Safety	
16.3.2.3.1.1	ECG Results SP-624 Subjects (Part 1)	Safety	
16.3.2.3.1.2	ECG Results Placebo Subjects (Part 2)	Safety	
16.3.2.3.2	ECG Results for Subjects Meeting at Least One QTcF Criterion	Safety	
16.3.2.3.3	ECG Results for Subjects Meeting at Least One Heart Rate Criterion	Safety	
16.3.3.1.1	Serum Chemistry Results SP-624 Subjects (Part 1)	Safety	
16.3.3.1.2	Serum Chemistry Results SP-624 Subjects (Part 2)	Safety	
16.3.3.1.3	Serum Chemistry Results Placebo Subjects (Part 1)	Safety	
16.3.3.1.4	Serum Chemistry Results Placebo Subjects (Part 2)	Safety	
16.3.3.2.1	Hematology Results SP-624 Subjects (Part 1)	Safety	

Listing Number	Title	Population	Included in Top Line Results Package
16.3.3.2.2	Hematology Results SP-624 Subjects (Part 2)	Safety	
16.3.3.2.3	Hematology Results Placebo Subjects (Part 1)	Safety	
16.3.3.2.4	Hematology Results Placebo Subjects (Part 2)	Safety	
16.3.3.3	Coagulation Results	Safety	
16.3.3.4.1	Urinalysis Results SP-624 Subjects	Safety	
16.3.3.4.2	Urinalysis Results Placebo Subjects	Safety	
16.3.3.5	Pregnancy Test Results	Safety	
16.3.3.6	Drug Screening Information	Safety	
16.3.3.7	Thyroid Function Test Results	Safety	
16.3.4.1	Physical Examination Results SP-624 Subjects (Part 1)	Safety	
16.3.4.2	Physical Examination Results Placebo Subjects (Part 2)	Safety	
16.3.5.1	C-SSRS Response Baseline/Screen Version	Safety	
16.3.5.2	C-SSRS Response Since the Last Visit Version by Visit	Safety	
16.3.6	PK Sampling Date and Time and Concentration Results	Safety	
16.3.7	Exploratory Biospecimen Sampling Dates and Times	Safety	

11 Table, Figure, and Listing Shells

To be provided in a separate document.