NCT02670551

Study ID: RGH-MD-54

Title: A Phase 3, Randomized, Double-Blind, Placebo-Controlled, Parallel-Group, Multicenter, Fixed-Dose Clinical Trial Evaluating The Efficacy, Safety And Tolerability Of Cariprazine In Patients With Bipolar I Depression

Statistical Analysis Plan Amendment 1 Date: 2017-Nov-29

1. Title Page

STATISTICAL ANALYSIS PLAN

A Phase 3, Randomized, Double-Blind, Placebo-Controlled, Parallel-Group, Multicenter, Fixed-Dose Clinical Trial Evaluating the Efficacy, Safety and Tolerability of Cariprazine in Patients with Bipolar I Depression

Original SAP Date: 2017-03-30

Amendment #1 Date: 2017-11-29

Protocol Number:	RGH-MD-54
Development Phase:	3
Product Name:	Cariprazine
Study Statistician:	[redacted]
Sponsor:	Forest Laboratories, LLC, an Allergan Affiliate

2. Table of Contents

2.1 List of Tables

2.2 List of Figures

3. List of Abbreviations and Definition of Terms

Table 3-1 Abbreviations and Definitions of Terms

Abbreviation/Term	Definition
AE	adverse event
AIMS	Abnormal Involuntary Movement Scale
ANCOVA	analysis of covariance
ATC	Anatomical Therapeutic Chemical
BARS	Barnes Akathisia Rating Scale
CFB	change from baseline
CGI-S	Clinical Global Impressions–Severity
C-SSRS	Columbia–Suicide Severity Rating Scale
eCRF	electronic case report form
ECG	electrocardiogram, electrocardiographic
ET	early termination
HAMD-17	17-item Hamilton Depression Rating Scale
HAM-A	Hamilton Anxiety Rating Scale
ITT	intent-to-treat
[redacted]	[redacted]
LS	least squares
MADRS	Montgomery-Åsberg Depression Rating Scale
MedDRA	Medication Dictionary for Regulatory Activities
MINI	Mini International Neuropsychiatric Interview
MMRM	mixed-effects model for repeated measures
PCS	potentially clinically significant
PK	pharmacokinetic
PID	participant identification
PT	preferred term
QIDS-SR_{16}	Quick Inventory of Depressive Symptomatology – Self-Report
QTc	QT interval corrected for heart rate
QTcB	QT interval corrected for heart rate using the Bazett formula ($QTcB = QT/(RR)^{1/2}$)
QTcF	QT interval corrected for heart rate using the Fridericia formula ($QTcF = QT/(RR)^{1/3}$)
SAE	serious adverse event
SAP	statistical analysis plan
SAS	Simpson-Angus Scale
SI	Le Système International d'Unités (International System of Units)
SOC	system organ class
TEAE	treatment-emergent adverse event
WHO	World Health Organization
YMRS	Young Mania Rating Scale

4. Introduction

This statistical analysis plan (SAP) details comprehensive, technical specifications of the statistical analyses of the efficacy and safety data outlined and/or specified in the final protocol of Study RGH-MD-54 and the most recent amendment (version #2). Specifications of tables, figures, and data listings are contained in a separate document.

This document is organized into 3 main sections:

1. Study overview
2. Statistical Methodology and Study Endpoints
3. Data Handling and Analysis Conventions

4.1 Study Design Summary

Methodology:

Multicenter, randomized, double-blind, placebo-controlled, parallel-group, fixed-dose study

Number of Patients:

480 planned (approximately 160 per treatment group)

Diagnosis and Main Criteria for Inclusion:

Male and female outpatients who are 18 to 65 years of age, meet Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) criteria for bipolar I disorder without psychotic features confirmed by the administration of the Mini International Neuropsychiatric Interview (MINI), with a current major depressive episode of at least 4 weeks and not exceeding 12 months in duration, a minimum score of 20 on the 17-item Hamilton Depression Rating Scale (HAMD-17), a minimum score of 2 on item 1 of the HAMD-17, and a minimum score of 4 on the Clinical Global Impressions-Severity Scale (CGI-S) at Visits 1 and 2.

Test Product, Dosage, and Mode of Administration:

Cariprazine 1.5 mg/day and 3.0 mg/day capsules, oral administration

Duration of Treatment:

A no-drug screening period of approximately 7-14 days, followed by 6 weeks of double-blind treatment and a 1-week, no investigational product safety follow-up period.

Reference Therapy, Dosage, and Mode of Administration:

Matching placebo capsules, oral administration

[redacted]

4.2 Study Objectives and Endpoints

Each study objective is presented with corresponding endpoint(s) below:

Table 4-1 Study Objectives and Endpoints

Objectives	Endpoints
• To evaluate the efficacy of cariprazine 1.5 mg/day and 3.0 mg/day relative to placebo in patients with bipolar I depression	**Primary Efficacy Endpoint** • Change from baseline to Week 6 in Montgomery-Åsberg Depression Rating Scale (MADRS) total score **Secondary Efficacy Endpoint** • Change from baseline to Week 6 in Clinical Global Impressions-Severity Scale (CGI-S)

Objectives	Endpoints
[redacted]	[redacted]
[redacted]	[redacted]
	[redacted]

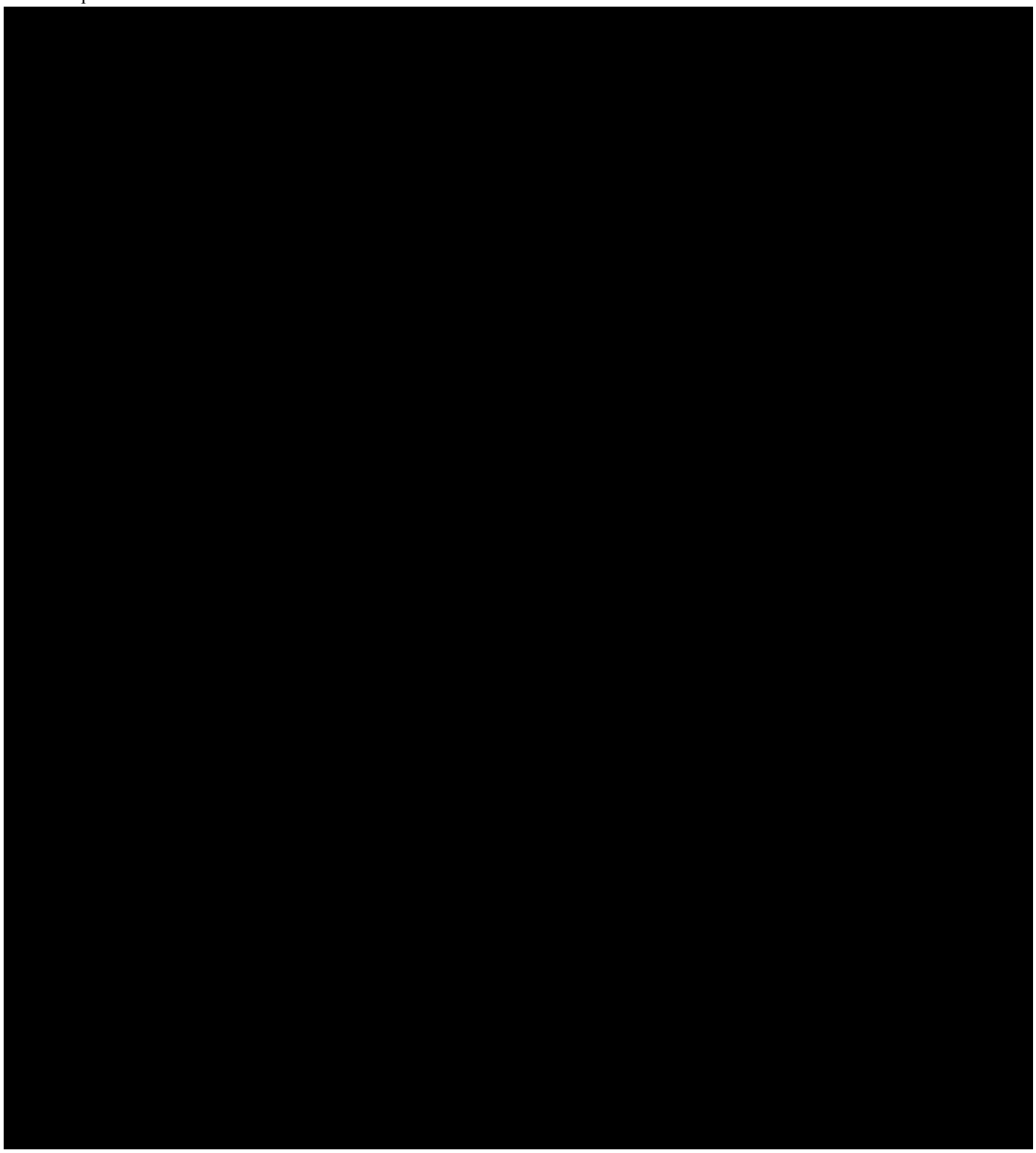
Table 4-2 Montgomery-Åsberg Depression Rating Scale (MADRS)

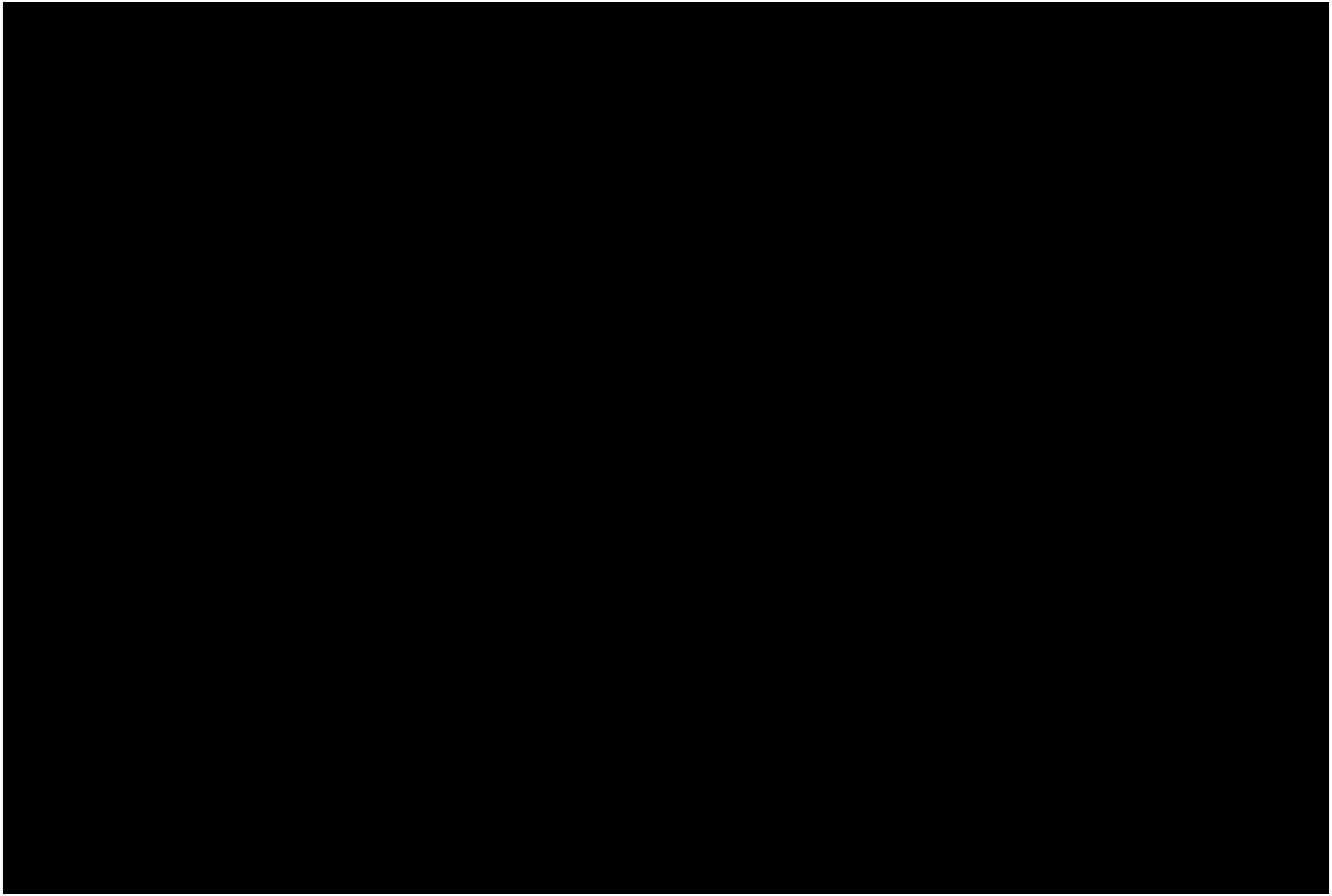
Item List[a]	Score Range
1 Apparent sadness 2 Reported sadness 3 Inner tension 4 Reduced sleep 5 Reduced appetite 6 Concentration difficulties 7 Lassitude 8 Inability to feel 9 Pessimistic thoughts 10 Suicidal thoughts	(Each item is rated on a scale from 0 to 6.)

[a] Scale for each item described in Appendix IV of the protocol.

Table 4-3 Clinical Global Impressions–Severity (CGI-S)

Grade	Description
1	Normal, not at all ill
2	Borderline ill
3	Mildly ill
4	Moderately ill
5	Markedly ill
6	Severely ill
7	Among the most extremely ill patients





5. Statistical Methodology and Study Endpoints

5.1 Statistical Methods Planned in the Protocol and Determination of Sample Size

This statistical analysis plan (SAP) will be approved prior to database lock. The SAP expands the statistical section of the protocol and contains a detailed description of methods to analyze data collected in the study. The text portion of the SAP will be included in the CSR report as Appendix 16.1.9.

5.1.1 Statistical and Analytical Plans

Statistical analyses will be conducted using SAS Version 9.3 or newer.

5.1.1.1 Common Conventions

5.1.1.1.1 Analysis Populations

The analysis populations will consist of participants as defined below:

Table 5-1 Analysis Populations

Population	Definition	Study Treatment
Screened	All patients who signed informed consent and received a PID number	—
Randomized	All patients in the Screened Population who were randomized to a treatment group in the study.	Randomized assignment
Safety	All patients in the Randomized Population who took at least 1 dose of double-blind investigational product	Randomized assignment
Intent-to-Treat (ITT)	All patients in the Safety Population who had at least 1 postbaseline assessment of the MADRS total score.	Randomized assignment

5.1.1.1.2 Study Treatments

The following treatment groups are defined for this study:

- Cariprazine 1.5 mg/day
- Cariprazine 3.0 mg/day
- Placebo

5.1.1.1.3 Statistical Methodology

The methodologies defined below apply as specified to individual endpoints defined in this SAP. When indicated, 95% two-sided confidence intervals and two-sided p-values will be presented. Individual endpoint analyses may specify exceptions (additions or deletions) to these general definitions.

Table 5-2 Statistical Methodology

Methodology	Description
Categorical counts	• Number of participants in individual categories ○ Participants with ≥ 1 qualifying event counted once per individual category

Categorical descriptives	• Number and percentage of participants in individual categories
	  ○ Participants with ≥ 1 qualifying event counted once per individual category
	  ○ N1 = number of patients with non-missing baseline
	  ○ For AE summary of sex-specific AEs, the percentage is based on the number of participants of the sex

PCS descriptives	• Number and percentage of participants meeting potentially clinically significant (PCS) criteria
	  ○ Participants with ≥ 1 qualifying event counted once per PCS category
	• Percentage denominator = number of participants with non-missing baseline and >=1 non-missing postbaseline assessment
	  ○ Unevaluable assessments considered missing
	• N1 = participants with non-missing values at both baseline and the specified postbaseline analysis visit

Shift analysis	• Number and percentage of participants in individual baseline and postbaseline categories • Percentage denominator = number of participants in individual baseline categories • N1 = participants with non-missing values at both baseline and the specified postbaseline analysis visit
Continuous descriptives	• N, mean, standard deviation (SD), median, minimum, maximum
CFB descriptives	• Continuous descriptives for baseline, postbaseline, and change from baseline (CFB) values
[redacted]	[redacted]

CFB MMRM	• Continuous descriptives and SE for baseline, postbaseline, and CFB values at each analysis visit
	• Estimates derived from mixed model for CFB value controlling for fixed factors (treatment group, pooled study center, and visit), covariates (baseline value), and interactions (treatment group by visit, baseline value by visit), with an unstructured covariance matrix (compound symmetry if convergence fails). The Kenward-Roger approximation will be used to estimate denominator degrees of freedom.
	  ○ LS means and standard errors
	  ○ LS mean differences, standard errors, and for comparisons of Cariprazine 1.5 mg/day and Cariprazine 3.0 mg/day vs Placebo
	  ○ P-values from contrast t-test comparing Cariprazine 1.5 mg/day and Cariprazine 3.0 mg/day vs Placebo (Note ***the nominal*** p-values for ***primary and secondary endpoints will be adjusted through*** the matched parallel gatekeeping procedure***, as described in Section 5.1.1.3.2.)***

Methodology	Description
[redacted]	[redacted]
[redacted]	[redacted]

CFB = change from baseline; ANCOVA = analysis of covariance; MMRM = mixed model for repeated measures;

Raw and derived data listings will be provided, and will be fully defined in the table, figure, and data listing specification document.

[redacted]

[redacted]

[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

5.1.1.1.5 Site Pooling

For efficacy analyses in which study center is a factor, a small center will be defined as a center with fewer than 2 patients in at least 1 treatment group in the ITT Population. Small centers will be pooled to form pseudo-centers so that each treatment group includes at least 2 ITT patients within the center. Pooling will be done within each country using the following algorithm:

1. Based on the number of ITT patients, small centers will be ordered from the largest to the smallest, and centers of the same size will be ordered from the largest center number to the smallest center number.

2. The pooling process starts with the largest small center from the top, which will be pooled with the smallest from the bottom until a non-small center is formed.
3. The process will be repeated using the small centers left out after the first pass.
4. If any centers are left out at the end of the process, they will be pooled with the smallest pseudo-center. If there is more than 1 smallest pseudo-center, the pseudo-center with the smallest center number will be selected. In case that the pseudo center formed by pooling all small centers is still a small center, it will be pooled with the smallest non-small center. If there is more than 1 smallest non-small center, the one with the smallest center number will be selected.

The pseudo-centers will be used for all efficacy analyses when the model is adjusted for study center.

5.1.1.2 Demographics

5.1.1.2.1 Analysis Populations

The distribution of participants within the analysis populations will be summarized as follows:

Table 5-4 Analysis Population Summaries

Population	Description	Timing	Methodology
Screened Population	Distribution overall and within Study Center in total	—	Categorical counts
Randomized, Safety, and ITT Populations	Distribution overall and within Study Center in total and by treatment group	—	Categorical counts

5.1.1.2.2 Participant Disposition

Participant disposition encompasses the distribution of participants who enter, complete, and discontinue each specified analysis period, along with eCRF-reported discontinuation reasons from each respective analysis period. Participant disposition will be summarized as follows:

Table 5-5 Participant Disposition Summaries

Parameter	Description	Timing	Methodology
Screening disposition	Distribution in the Screened Population overall and within Study Center in total	Screening Period	Categorical descriptives
Double-blind disposition	Distribution in the Randomized, Safety and ITT Population overall and within Study Center in total and by treatment group	Double-blind Treatment Period	Categorical descriptives
Double-blind disposition	Distribution by reason in the Safety Population in total and by treatment group	Double-blind Treatment Period	Categorical descriptives

5.1.1.2.3 Protocol Deviations

Protocol deviations will be defined in Protocol Deviation Requirement Specification, including importance classification. Protocol deviations will be summarized as ***follows for the Randomized Population***:

Table 5-6 Protocol Deviation Summary

Parameter	Description	Timing	Methodology
Important protocol deviations[1]	Distribution in the ***Randomized*** Population in total and by treatment group	—	Categorical descriptives

[1]Protocol deviations will be listed.

5.1.1.2.4 Demographics

Demographics will be summarized in total and by treatment group for the ITT and Safety Populations, as follows:

Table 5-7 Demographic and Baseline Characteristics Summaries

Parameter	Description	Timing	Methodology
Age	Age (years) relative to informed consent date	Informed consent	Continuous descriptives
Age group	• < 20 years • ≥ 20 to < 30 years • ≥ 30 to < 40 years • ≥ 40 to < 50 years • ≥ 50 to < 60 years • ≥ 60 years	Informed consent	Categorical descriptives

Sex, race, and ethnicity	• eCRF categories
	• Race group
	  ○ White
	  ○ Non-white
		Screening Period	Categorical descriptives

5.1.1.2.5 Baseline Characteristics

Baseline characteristics will be summarized in total and by treatment group for the ITT and Safety Populations, as follows:

Table 5-8 Demographic and Baseline Characteristics Summaries

Parameter	Description	Timing	Methodology
Baseline characteristics	• Height (m) • Weight (kg) • Body mass index (BMI) • Weight (kg) / height (m^2)	Latest assessment in Screening Period	Continuous descriptives
Baseline efficacy	Endpoints and timing fully described in Section 5.1.1.3 • MADRS • CGI-S • [redacted] • Summary for ITT Population only	—	Continuous descriptives

5.1.1.2.6 Medical History

Medical history, encompassing abnormalities and surgeries reported as occurring before the Screening Visit, will be coded using the Medical Dictionary for Regulatory Activities (MedDRA), version 19.1 or newer. Unique participants who report medical history events will be summarized by MedDRA system organ class (SOC) and preferred term (PT) in total and by treatment group for the Safety Population as follows:

Table 5-9 Medical History Summary

Parameter	Description	Timing	Methodology
Medical and Surgical history	Abnormalities and surgeries occurring before the Screening Visit	Screening Period	Categorical descriptives
Bipolar History	Bipolar disorder history occurring before the Screening Visit	Screening Period	Categorical descriptives
Other Psychiatric Disorder Medical History	Other psychiatric history occurring before the Screening Visit	Screening Period	Categorical descriptives
Non-Drug Psychiatric Treatment History	Non-drug psychiatric treatments taken before the Screening Visit	Screening Period	Categorical descriptives
History of Ocular Events of Special Interest	Abnormalities of ocular events occurring before the Screening Visit	Screening Period	Categorical descriptives
Ophthalmic Medical and Surgical History	Ophthalmic medical and surgical history occurring before the Screening Visit	Screening Period	Categorical descriptives

SOCs will be sorted alphabetically; PTs will be sorted in descending frequency in the highest dose group.

5.1.1.2.7 Prior and Concomitant Medications

Medications will be coded using the World Health Organization (WHO) Drug Dictionary, version WHODRUG DDE + HD 3Q2016 or newer. Unique participants who reported medications will be summarized by Anatomical Therapeutic Chemical (ATC) 4 class and PT in total and by treatment group for the Safety Population as described in the table below. Additionally, the Concomitant Rescue Medications will be summarized, for each medication, with the mean daily dose, in total and by treatment group for the Safety Population.

Table 5-10 Medication Summaries

Parameter	Description	Timing	Methodology
Prior medications[1]	Medication taken before the date of the first dose of double-blind investigational product.	Screening Period	Categorical descriptives
Concomitant medications[1]	Medication taken on or after the date of the first dose of double-blind investigational product.	Double-blind Treatment Period	Categorical descriptives
Concomitant Medication of Interest- Summary of Rescue Medication Mean Daily Dose[1]	Rescue Medication taken on or after the date of the first dose of double-blind investigational product.	Double-blind Treatment Period	Continuous descriptives

ATC4 classes will be sorted alphabetically; PTs will be sorted in descending frequency in the highest dose group.
[1]Participant data will be listed.

5.1.1.3 Efficacy and Pharmacokinetic Analyses

Efficacy analyses will be based on the ITT Population. Pharmacokinetic endpoints and analyses will not be described in this SAP.

The following efficacy assessments and terms are defined:

Table 5-11 Efficacy Assessments

Assessment/Term	Description
Montgomery-Åsberg Depression Rating Scale (MADRS)	10-Item, clinician-rated scale to evaluate the patient's depressive symptomatology. Each item is scored on a 7-point scale, from 0 to 6 with a score of 0 reflecting no symptoms and a score of 6 reflecting symptoms of maximum severity. The MADRS total score is the sum of the 10 items from the MADRS ***at each visit if no item has missing score. If the number of missing items at a visit is 2 or less, the MADRS total score will be calculated as (sum of nonmissing items) × (total number of items) / (number of nonmissing items).*** If more than 2 items of the MADRS are missing, then the total score will be set to missing.
Clinical Global Impressions–Severity (CGI-S)	Clinician-rated scale that measures the overall severity of a patient's illness in comparison with the severity of other patients the physician has observed. The patient is rated on a scale from 1 to 7 with 1 indicating a "normal state" and 7 indicating "among the most extremely ill patients.".
[redacted]	[redacted]

Assessment/Term	Description
	[redacted]

Baseline assessments for applicable efficacy endpoints defined as follows:

Table 5-12 Efficacy Endpoint Baseline Definitions

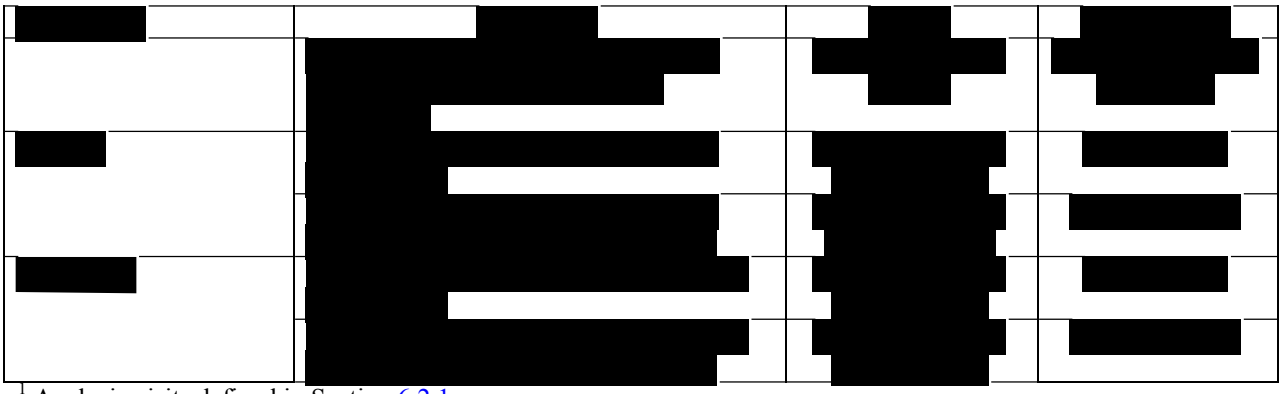
Endpoint	Description	Timing
Clinical Global Impressions–Severity (CGI-S) and the total scores of the following: Montgomery-Åsberg Depression Rating Scale (MADRS), [redacted]	The baseline for each specific efficacy endpoint is defined as the value recorded at Visit 2/Week 0. If this value is not available, the last available value before the first dose of double-blind investigational product will be used as the baseline value.	Treatment Day 1. If not available, then the last available value before the first dose from Day -14 to Day 1

5.1.1.3.1 Efficacy Endpoint

The following efficacy variables will be analyzed as described below.

Table 5-13 Efficacy Analyses

Assessment	Parameter	Timing	Methodology
MADRS	Change from baseline in MADRS total score to Week 6	Treatment Period: Week 6	CFB MMRM
	[redacted]	[redacted]	[redacted]
CGI-S	Change from baseline in CGI-S score by visit	Treatment Period: Weeks 1, 2, 4, and 6	CFB MMRM
	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]



[1] Analysis visits defined in Section 6.2.1.

5.1.1.3.2 Multiple Comparisons Procedure for Primary and Secondary Endpoints

Procedure:

To control the overall type I error rate for multiple comparisons of 2 active doses versus placebo at Week 6 for the primary and secondary efficacy parameters, the parallel gatekeeping procedure will be implemented (Chen et al, 2005).

Hypotheses:

The 4 null hypotheses will be grouped into 2 families denoted by F1 and F2, with F1 consisting of 2 null hypotheses for comparisons of the 2 active doses versus placebo in change from baseline to Week 6 in MADRS total score ***(i.e., $F_1 = \{H_{11}, H_{12}\}$)***, and F2 consisting of 2 null hypotheses for comparisons of the 2 active doses versus placebo in change from baseline to Week 6 in CGI-S score ***(i.e., $F_2 = \{H_{21}, H_{22}\}$)***. Family F1 will serve as the parallel gatekeeper for F2.

Testing:

The matched gatekeeping procedure utilizes the special logical relationship between the primary and the secondary parameters to enhance the power of statistical testing. The secondary efficacy parameter will be tested at a specific dose only if the corresponding primary efficacy parameter is statistically significant. The Simes test will be used to derive the local p-values for the interaction hypotheses. The adjusted p-values for the 4 elementary hypotheses will be calculated based on the closed testing principle. Statistical significance will be determined by comparing the adjusted p-values to $\alpha = 0.050$.

Weights for Intersection Tests in the Matched Parallel Gatekeeping Procedure

Intersection hypothesis	***Weights***			
	H_{11}	H_{12}	H_{21}	H_{22}
$H_{11} \cap H_{12} \cap H_{21} \cap H_{22}$	0.5	0.5	0	0
$H_{11} \cap H_{12} \cap H_{21}$	0.5	0.5	0	0
$H_{11} \cap H_{12} \cap H_{22}$	0.5	0.5	0	0
$H_{11} \cap H_{12}$	0.5	0.5	0	0
$H_{11} \cap H_{21} \cap H_{22}$	0.5	0	0	0.5
$H_{11} \cap H_{21}$	1	0	0	0
$H_{11} \cap H_{22}$	0.5	0	0	0.5
H_{11}	1	0	0	0
$H_{12} \cap H_{21} \cap H_{22}$	0	0.5	0.5	0
$H_{12} \cap H_{21}$	0	0.5	0.5	0
$H_{12} \cap H_{22}$	0	1	0	0
H_{12}	0	1	0	0
$H_{21} \cap H_{22}$	0	0	0.5	0.5
H_{21}	0	0	1	0
H_{22}	0	0	0	1

5.1.1.4 Safety Analyses

Safety analyses will be based on the Safety Population.

Baseline assessments for applicable safety endpoints defined as follows:

[redacted]

5.1.1.4.1 Study Treatment Exposure and Compliance

Study treatment exposure and compliance will be summarized by treatment group for the Safety Population as follows:

Table 5-15 Study Treatment Summaries

Parameter	Description	Timing	Methodology
Study treatment exposure (days)	Treatment end date - treatment start date + 1	Treatment Period	Continuous descriptives
Patient-years of study treatment exposure (years)	Sum over participants of study treatment exposure	Treatment Period	—

Parameter	Description	Timing	Methodology
Study treatment compliance (%)	Summary by visit interval and overall $100 \times \frac{\text{Number of capsules actually taken}}{\text{Number of capsules expected to be taken}}$ • The total number of capsules actually taken during a specific period is calculated as the sum of the number of capsules taken during that interval. • The number of capsules expected to be taken for a specific interval will be calculated by multiplying the number of days in that period by the number of capsules to be taken per day.	Week 0 (Baseline)-Week 1 visit (not inclusive), Week 1 visit-Week 2 visit (not inclusive), Week 2 visit-Week 4 visit (not inclusive), Week 4 visit-Week 6 visit (not inclusive), and Overall Treatment Period	Continuous descriptives

5.1.1.4.2 Adverse Events

The following adverse event (AE) terms are defined:

Table 5-16 AE Terms

Term	Description
Treatment-emergent	An event that initially occurs or increases in severity on or after the treatment start date, where: • Treatment start date ≤ event start date ≤ treatment end date + 30
Newly-emergent	An AE occurring during the safety follow-up period will be considered a newly emergent adverse event if the AE was not present before the start of the safety follow-up period or was present before the start of the safety follow-up period and increased in severity during the safety follow-up period.
On-therapy	An event where: • Treatment start date ≤ event start date ≤ treatment end date + 30

AEs, encompassing abnormalities and surgeries reported as occurring after the Screening Visit, will be coded using MedDRA version 19.1 or newer. Unique patients reporting AEs in the following AE categories will be summarized by treatment group for the Safety Population as follows:

Table 5-17 AE Summaries

Parameter	Description	Timing	Methodology
Overall summary	Overall summary only for the following categories: • Deaths • Treatment-emergent AEs (TEAEs) • On-therapy Serious adverse events (SAEs) • AEs leading to discontinuation	Double-blind Treatment Period	Categorical descriptives
	Overall summary only for the following categories: • Deaths • Serious adverse events (SAEs) • Newly emergent AEs	During the Safety Follow-up Period and within 30 days after the last dose of double-blind treatment	Categorical descriptives
TEAEs	Overall summary and by SOC and PT	Double-blind Treatment Period	Categorical descriptives
NEAEs	Overall summary and by SOC and PT	During the Safety Follow-up Period and within 30 days after double-blind treatment	Categorical descriptives
Common TEAEs	Summary by SOC, PT, sorted by decreasing frequency for the test treatment • Includes TEAEs occurring in ≥ 2.0% of participants in any treatment group. Event rates will not be rounded prior to comparison with the 2.0% cutpoint. Summary by PT, sorted by decreasing frequency for the test treatment	Double-blind Treatment Period	Categorical descriptives
TEAEs by severity	Overall summary and by SOC, PT, and severity • Participants categorized overall and within each SOC and PT for the most sever occurrence	Double-blind Treatment Period	Categorical descriptives
TEAEs by causality	Overall summary and by SOC, PT, and causality	Double-blind Treatment Period	Categorical descriptives
On-therapy SAEs[1]	Overall summary and by SOC and PT	Double-blind Treatment Period and within 30 days after the last dose of double-blind treatment;	Categorical descriptives
On-therapy SAEs[1]	Overall summary and by SOC and PT	During the Safety Follow-up Period and within 30 days after the last dose of double-blind treatment	Categorical descriptives

Parameter	Description	Timing	Methodology
AEs leading to discontinuation[1]	Overall summary and by SOC and PT, sorted by decreasing frequency for the highest dose test treatment	Double-blind Treatment Period	Categorical descriptives
AE of interest: Extra-Pyramidal Symptoms (EPS) related TEAEs[1]	Overall summary and by PT. Include the below PT: Akathisia, Bradykinesia, Cogwheel rigidity, Drooling, Dyskinesia, Dystonia, Extrapyramidal disorder, Hypokinesia, Masked facies, Muscle rigidity, Muscle tightness, Musculoskeletal stiffness, Oculogyric crisis, Oromandibular dystonia, Parkinsonism, Restlessness, Salivary hypersecretion, Tardive dyskinesia, Tongue spasm, Tremor, Trismus	Double-blind Treatment Period	Categorical descriptives
AE of interest: Ocular events[1]	Overall summary and by Category and PT • Includes Categories and PTs, defined in Section 6.6.1.2	Double-blind Treatment Period	Categorical descriptives

[1] Patients who report ≥ 1 AE in the AE category and all AEs for those patients will be listed.
SOCs will be sorted alphabetically; PTs will be sorted in descending frequency in the highest dose group.

[redacted]

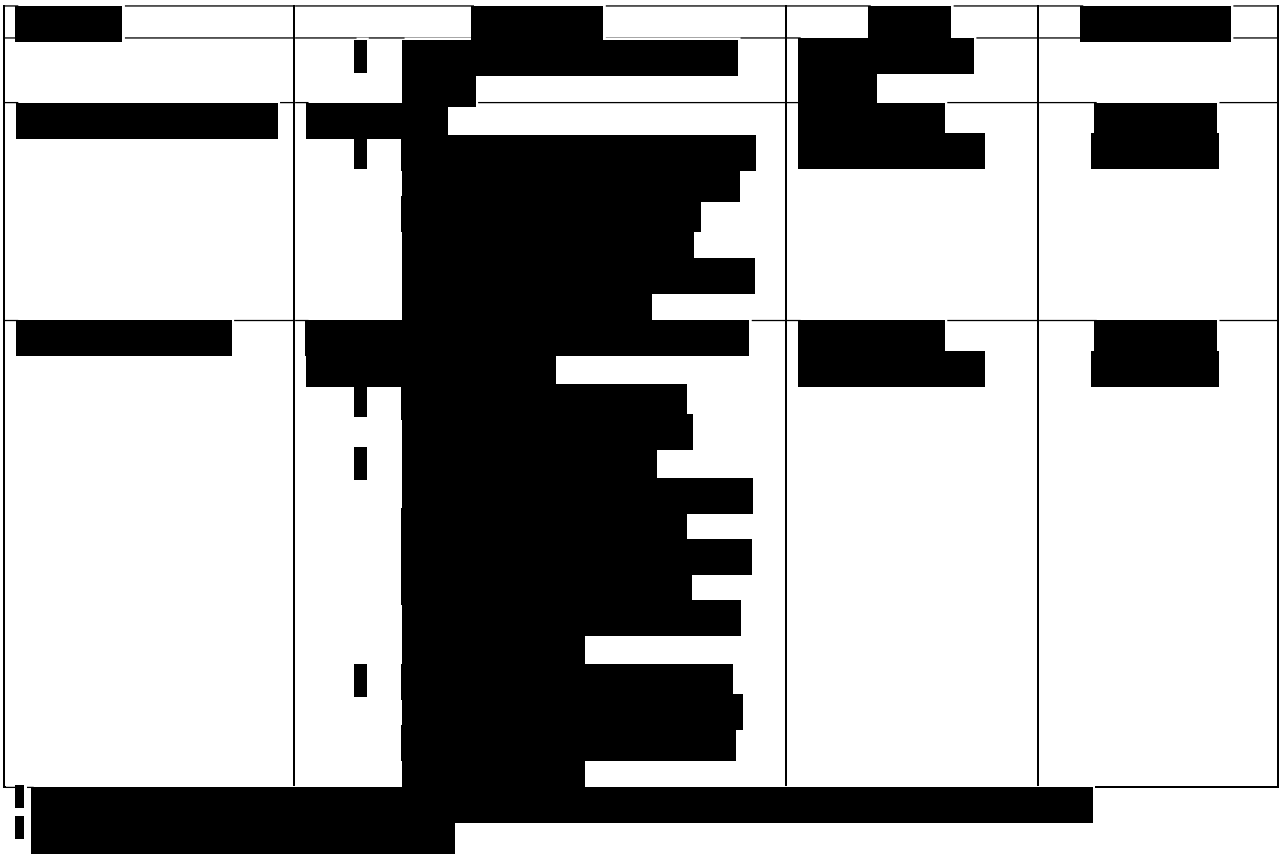
Endpoint	Description	Timing	Methodology
[redacted]	[redacted]	[redacted]	[redacted]

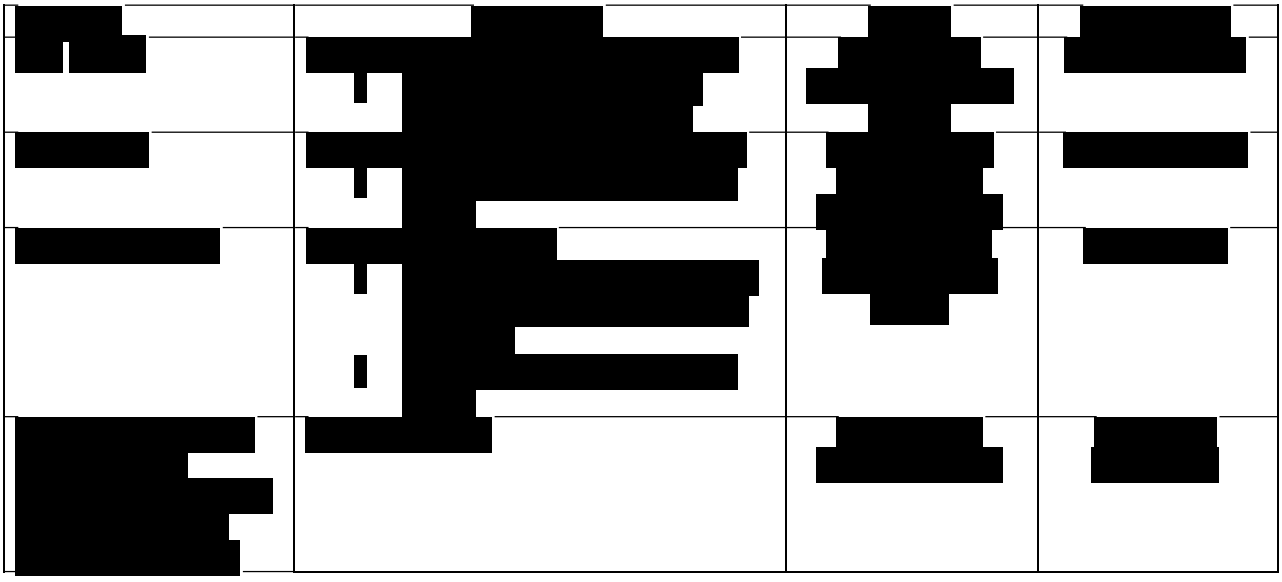
[redacted]

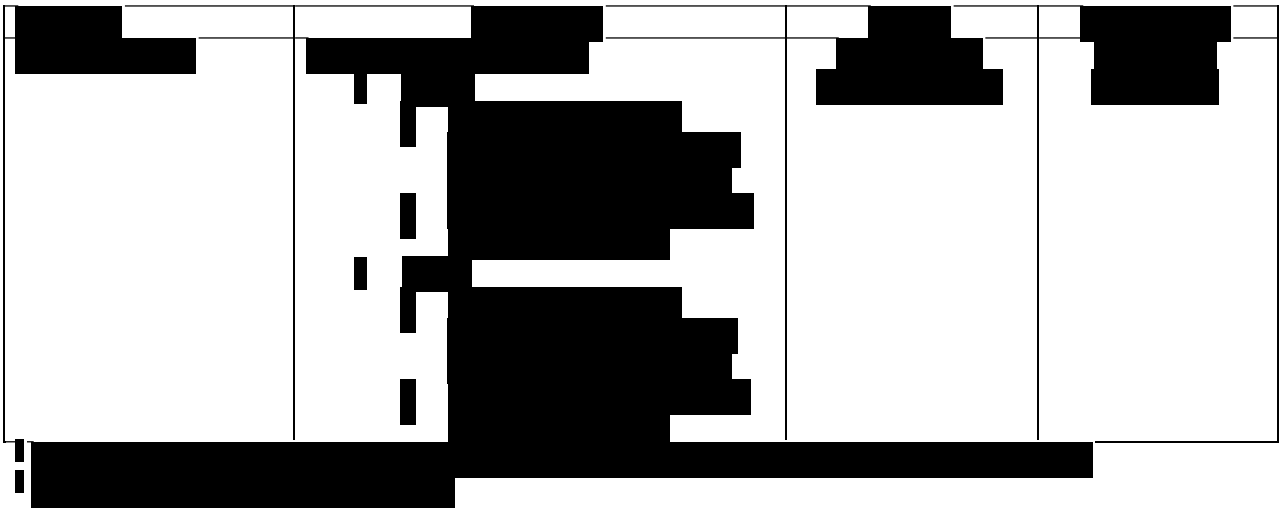
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

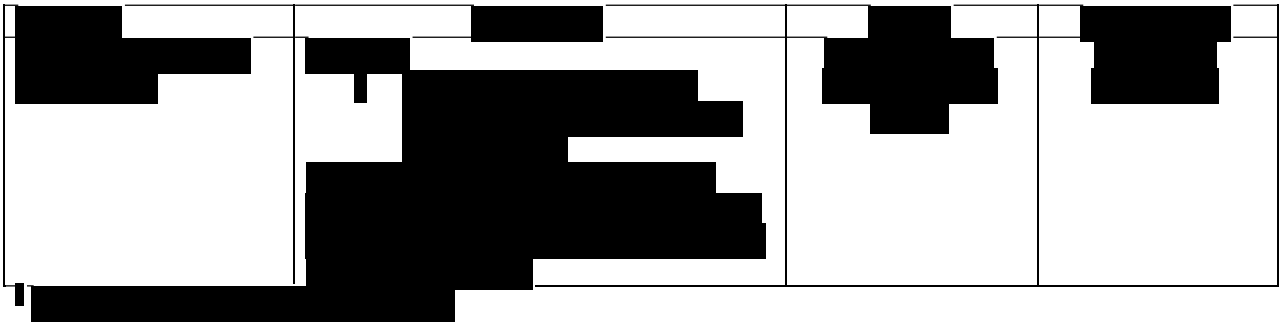
[redacted]

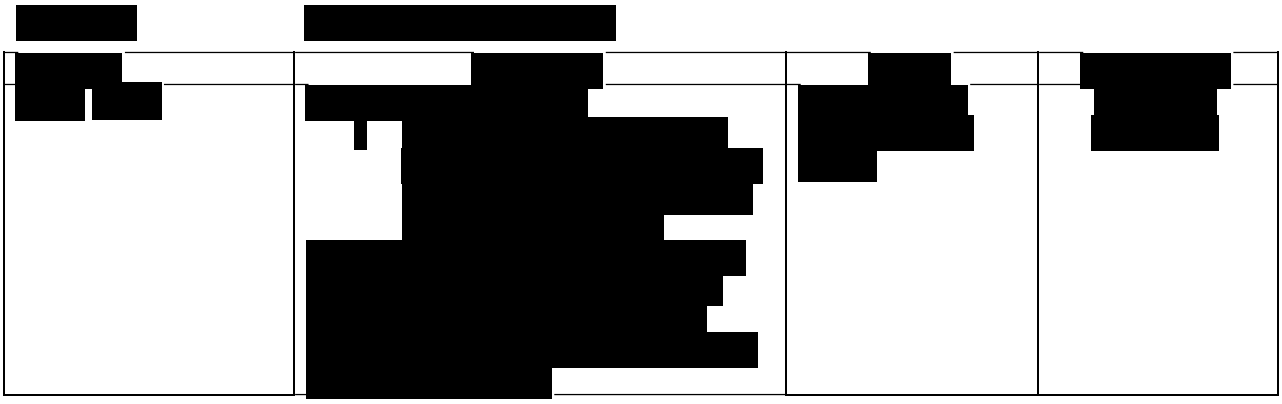
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]











5.1.1.5 Subgroup Analyses

Not applicable.

5.1.1.6 Interim Analyses

No interim analysis is planned for this study.

5.1.2 Determination of Sample Size

The primary efficacy endpoint is the change from baseline at Week 6 in MADRS total score. The secondary endpoint is the change from baseline at Week 6 in CGI-S score. For the comparison of

the primary endpoint, the sample size of 160 subjects per arm will provide approximately 82% statistical power to show statistically significantly higher effect in each dose of cariprazine versus placebo. The study has statistical power of 90% to show at least one of the two cariprazine doses is statistically significantly more efficacious than placebo in the primary endpoint. These calculations assumed an effect size of 0.36 (treatment group difference relative to SD). For a given successful (statistically significant) dose versus placebo with respect to the primary endpoint, its probability in detecting an effect size of 0.36 as statistically significant versus placebo with respect to the secondary endpoint is 88%. All statistical powers presented in this section were calculated adjusting for multiple comparisons using matched parallel gatekeeping procedure with the family-wise Type I error rate being controlled at a 0.05 level (2-sided). The dropout rate is assumed to be 22% at Week 6. Within-person correlation for both primary and secondary endpoints is assumed to be 0.6, as well as correlation between the two endpoints (primary and secondary) to be 0.6. Assumptions of effect sizes, correlation coefficients, and drop-out rate are based on RGH-MD-56 study.

5.2 Changes in the Conduct of the Study or Planned Analyses

This analysis plan is consistent with the methods stated in the study protocol. Changes in the conduct of the study or planned analyses, should they happen, will be documented in the clinical study report.

5.2.1 Changes in the Conduct of the Study

Not applicable.

5.2.2 Changes to Analyses Prior to Database Lock

Not applicable.

[redacted] [redacted]

[redacted] [redacted]

[redacted] [redacted]

[redacted]

[redacted] [redacted]

[redacted]	[redacted]
[redacted]	[redacted]

[redacted] [redacted]

[redacted]

[redacted] [redacted]

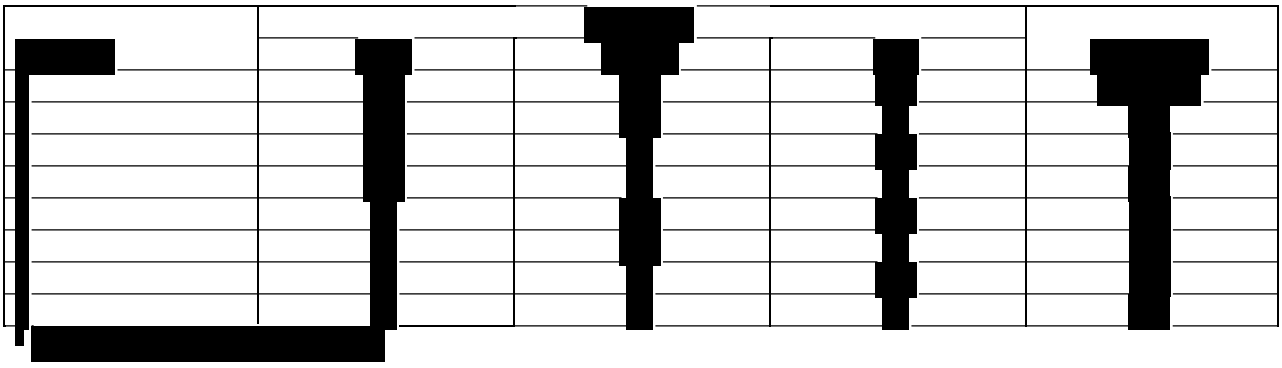
[redacted] [redacted]

[redacted]

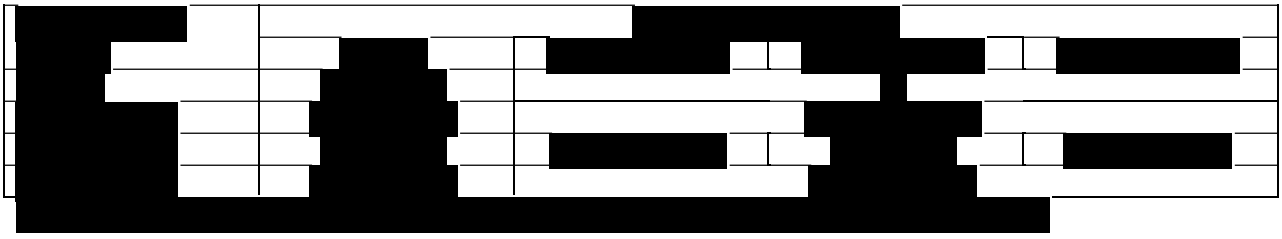
[redacted] [redacted]

[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted] eek 6/Early Termination visit]

[redacted]



[redacted]



[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

6.6.1.2 AE Group of Interest

The following PTs are categorized as AE groups of special interest for Ocular Events:

Table 6-6 AE of Special Interest Terms: Ocular Events

Category of Ocular Events	Verbatim	Preferred Terms
Cataract Category	Cataract	Atopic cataract, Cataract, Cataract cortical, Cataract diabetic, Cataract nuclear, Cataract subcapsular, Toxic cataract
	Lens or lenticular abnormality	Acquired lenticonus, Aphakia, Lens disorder
	Lens opacity, opacification or opalescence	Lens discolouration, Lenticular opacities, Posterior capsule opacification
Visual Acuity Change Category	Blindness	Amaurosis, Amaurosis fugax Blindness, Blindness day, Blindness transient, Blindness unilateral, Diabetic blindness, Sudden visual loss
	Night blindness	Delayed dark adaptation, Night blindness
	Visual acuity decreased, abnormality or change	Refraction disorder, Visual acuity reduced, Visual acuity reduced transiently
	Vision decreased, abnormality or change	Accommodation disorder, Altered visual depth perception, Aniseikonia, Anisometropia, Antimetropia, Astigmatism, Hemianopia, Hemianopia heteronymous, Hypermetropia, Myopia, Pathologic myopia, Presbyopia, Pseudomyopia, Quadranopia, Tunnel vision, Visual field defect, Visual impairment
	Visual acuity test abnormality or change	Visual acuity tests abnormal
Retinal Impairment Category	Retinal degeneration, abnormality or change	Pars plana cyst, Retinal cyst, Retinal degeneration, Retinal depigmentation, Retinal deposits, Retinal detachment, Retinal disorder, Retinal dystrophy, Retinal fibrosis, Retinal infiltrates, Retinal pallor, Retinal pigmentation, Retinal pigment epithelial tear, Retinal pigment epitheliopathy, Retinal tear, Retinal thickening Retinal toxicity, Retinoschisis, Subretinal fibrosis, Subretinal fluid
	Macular degeneration, abnormality or change	Age-related macular degeneration, Dry age-related macular degeneration, Macular cyst, Macular degeneration, Macular hole, Macular opacity, Macular detachment, Macular rupture, Macular fibrosis, Macular pigmentation, Macular rupture, Maculopathy, Neovascular age-related macular degeneration, Vitreomacular interface abnormal
	Optic nerve degeneration, abnormality or change	Myopic disc, Optic atrophy, Optic discs blurred, Optic disc disorder, Optic disc drusen, Optic nerve cupping, Optic nerve cup/disc ratio decreased, Optic nerve cup/disc ratio increased, Optic nerve disorder, Papilloedema, Pseudopapilloedema
	Retinal pigment epithelium detachment, abnormality or change	Chorioretinal atrophy, Chorioretinal disorder, Chorioretinopathy, Detachment of macular retinal pigment epithelium, Detachment of retinal pigment epithelium, Retinal dystrophy
Color Vision Change Category	Color vision decrease, abnormality or change	Chloropsia, Chromatopsia, Colour blindness, Colour blindness acquired, Colour vision tests abnormal, Colour vision tests abnormal blue-yellow, Colour vision tests abnormal red-green, Cyanopsia, Erythropsia, Xanthopsia

[redacted] [redacted]

[redacted] [redacted]

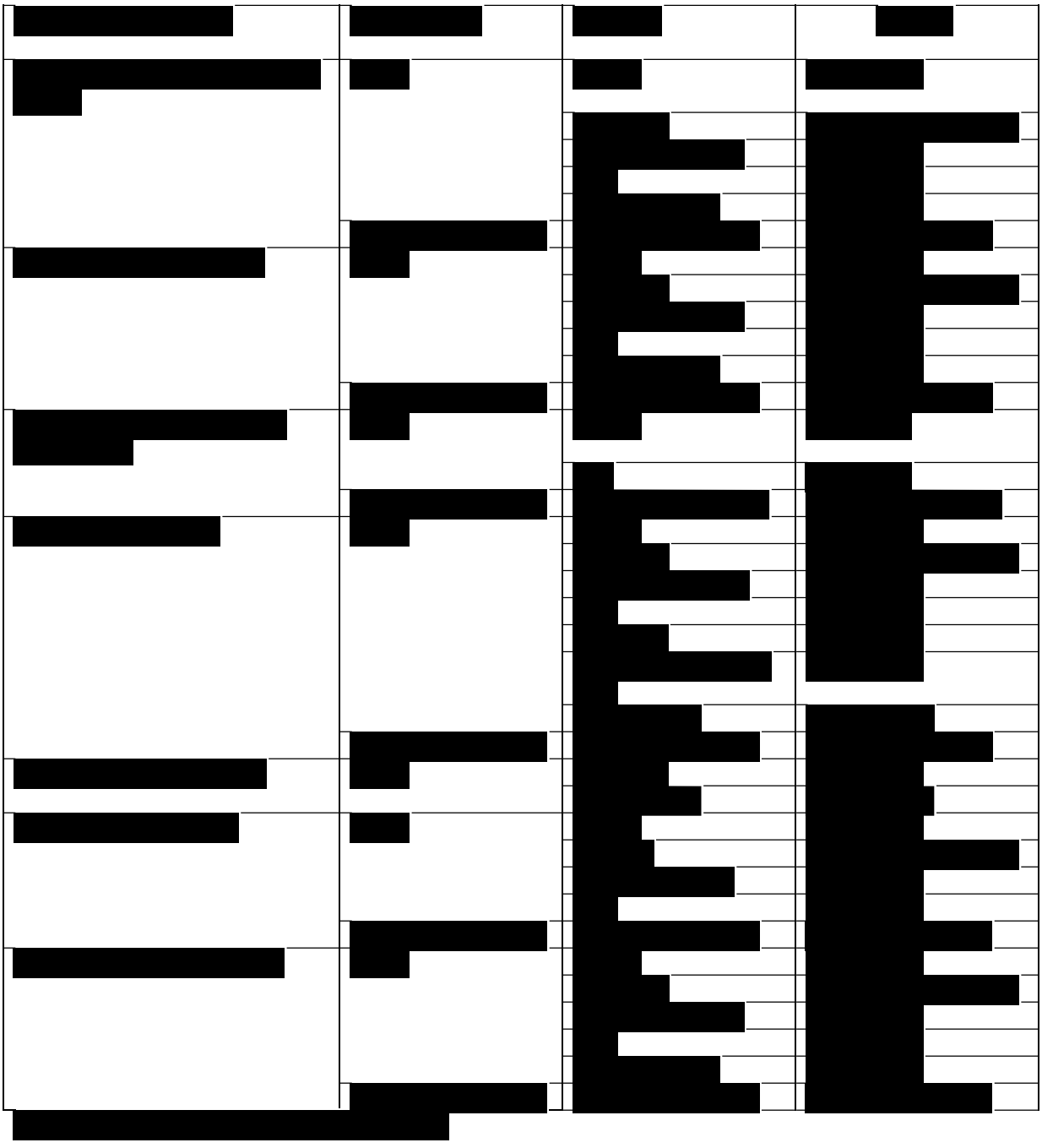
[redacted]

[redacted] [redacted]

[redacted]	[redacted]	[redacted]	[redacted]
[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]			

[redacted]

[redacted]



[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted] [redacted]

[redacted]

[redacted] [redacted]

	[redacted]	
[redacted]	[redacted]	[redacted]

[redacted]

[redacted] [redacted]

[redacted]

[redacted] [redacted]

[redacted]		[redacted]
[redacted]	[redacted]	[redacted]

[redacted]

[redacted] [redacted]

[redacted] [redacted]

[redacted]

[redacted] [redacted]

[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

[redacted]

[redacted] [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

7. References

Chen X et al. The application of enhanced parallel gatekeeping strategies. *Statistics in Medicine,* 2005; 24 :1385-1397.

Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry 1960; 23:56–62.

Kenward MG, Molenberghs G, Thijs H. Pattern-mixture models with proper time dependence. *Biometrika*, 2003; 90(1): 53-71.

Rubin DB. Multiple Imputation for Nonresponse in Surveys. J. Wiley & Sons, New York; 1987.

Rush JA, Trivedi MH, Ibrahim HM, et al. The 16-item Quick Inventory of Depressive Symptomatology (QIDS), Clinician Rating (QIDS-C), and Self-Report (QIDS-SR): A psychometric evaluation in patients with chronic major depression. Biol Psychiatry 2003;573-83.

SAS Institute Inc. SAS/STAT User's Guide, Version 9.1. Cary, NC: SAS Institute Inc; 2004.

Wang C, Daniels MJ. A note on MAR, identifying restrictions, model comparison, and sensitivity analysis in pattern mixture models with and without covariates for incomplete data. *Biometrics,* 2011; 67(3):810-818.

STATISTICAL ANALYSIS PLAN AMENDMENT
SUMMARY OF CHANGES

A Phase 3, Randomized, Double-Blind, Placebo-Controlled, Parallel-Group, Multicenter, Fixed-Dose Clinical Trial Evaluating the Efficacy, Safety and Tolerability of Cariprazine in Patients with Bipolar I Depression

Original SAP Date: 2017-03-30

Amendment #1: 2017-11-29

Protocol Number:	RGH-MD-54
Development Phase:	3
Product Name:	Cariprazine
Study Statistician:	[redacted]
Sponsor:	Forest Laboratories, LLC, an Allergan Affiliate

1. Introduction

Amendment #1 specifies the following changes to the original statistical analysis plan (SAP) for Study RGH-MD-54, dated March 30, 2017:

- add a section for missing/imcomplete onset date of current episode of bipolar 1 disorder
- make revision to
  - update population for protocol deviation analysis
  - update imputation rules for missing items in efficacy parameter
  - update visits for analysis per design
  - include the weights for intersection tests in multiple comparison procedure
  - update analysis window for safety parameters to be consistent with previous studies
  - update lab PCS critiera
  - update lab parameters included in the summary analysis
  - update lab character value coding convention
  - make editorial updates that are necessary

These changes to the SAP are also reflected in the amended SAP. The page and section numbers in the headings of this amendment are those of the original SAP, dated March 30, 2017.

2. Global Changes

None

[redacted]

4. Deleted Sections

None

5. Revisions

5.1 SECTION 5.1.1.1.3, Statistical Methodology (Page 15)

Rational: This section has been amended to include details needed editorial updates of the MMRM model.

This section now reads as follows:

The methodologies defined below apply as specified to individual endpoints defined in this SAP. When indicated, 95~~.0~~% two-sided confidence intervals and two-sided p-values will be presented. Individual endpoint analyses may specify exceptions (additions or deletions) to these general definitions.

Table 5-2 Statistical Methodology

Methodology	Description
Categorical counts	• Number of participants in individual categories ○ Participants with ≥ 1 qualifying event counted once per individual category

Categorical descriptives	• Number and percentage of participants in individual categories
	  ○ Participants with ≥ 1 qualifying event counted once per individual category
	  ○ N1 = number of patients with non-missing baseline
	  ○ For AE summary of sex-specific AEs, the percentage is based on the number of participants of the sex

PCS descriptives	• Number and percentage of participants meeting potentially clinically significant (PCS) criteria
	  ○ Participants with ≥ 1 qualifying event counted once per PCS category
	• Percentage denominator = number of participants with non-missing baseline and >=1 non-missing postbaseline assessment
	  ○ Unevaluable assessments considered missing
	• N1 = participants with non-missing values at both baseline and the specified postbaseline analysis visit

Shift analysis	• Number and percentage of participants in individual baseline and postbaseline categories • Percentage denominator = number of participants in individual baseline categories • N1 = participants with non-missing values at both baseline and the specified postbaseline analysis visit
Continuous descriptives	• N, mean, standard deviation (SD), median, minimum, maximum
CFB descriptives	• Continuous descriptives for baseline, postbaseline, and change from baseline (CFB) values
[redacted]	[redacted]

Methodology	Description
CFB MMRM	• Continuous descriptives and SE for baseline, postbaseline, and CFB values at each analysis visit • Estimates derived from mixed model for CFB value controlling for fixed factors (treatment group, pooled study center, and visit), covariates (baseline value), and interactions (treatment group by visit, baseline value by visit), with an unstructured covariance matrix (compound symmetry if convergence fails). The Kenward-Roger approximation will be used to estimate denominator degrees of freedom. o LS means and standard errors o LS mean differences, standard errors, and for comparisons of Cariprazine 1.5 mg/day and Cariprazine 3.0 mg/day vs Placebo o P-values from contrast t-test comparing Cariprazine 1.5 mg/day and Cariprazine 3.0 mg/day vs Placebo (Note that the nominal p-values for ~~endpoints in~~ primary and secondary endpoints will be adjusted through the matched parallel gatekeeping procedure, as described in Section 5.1.1.3.2~~will be adjusted~~.)
[redacted]	[redacted]
[redacted]	[redacted]

CFB = change from baseline; ANCOVA = analysis of covariance; MMRM = mixed model for repeated measures;

Raw and derived data listings will be provided, and will be fully defined in the table, figure, and data listing specification document.

5.2 SECTION 5.1.1.2.2, Participant Disposition (Page 19)

Rational: This section has been amended to remove the disposition analysis for Randomized Population.

This section now reads as follows:

Participant disposition encompasses the distribution of participants who enter, complete, and discontinue each specified analysis period, along with eCRF-reported discontinuation reasons from each respective analysis period. Participant disposition will be summarized as follows:

Table 5-5 Participant Disposition Summaries

Parameter	Description	Timing	Methodology
Screening disposition	Distribution in the Screened Population overall and within Study Center in total	Screening Period	Categorical descriptives
Double-blind disposition	Distribution in the Randomized, Safety and ITT Population overall and within Study Center in total and by treatment group	Double-blind Treatment Period	Categorical descriptives
Double-blind disposition	Distribution by reason in the ~~Randomized and~~ Safety Population in total and by treatment group	Double-blind Treatment Period	Categorical descriptives

5.3 SECTION 5.1.1.2.3, Protocol Deviations (Page 20)

Rational: This section has been amended to use Randomized Population for protocol deviation analysis.

This section now reads as follows:

Protocol deviations will be defined in Protocol Deviation Requirement Specification, including importance classification. Protocol deviations will be summarized as follows for the Randomized Population:

Table 5-6 Protocol Deviation Summary

Parameter	Description	Timing	Methodology
Important protocol deviations[1]	Distribution in the ~~ITT~~ Randomized Population in total and by treatment group	—	Categorical descriptives

[1]Protocol deviations will be listed.

5.4 SECTION 5.1.1.3, Efficacy and Pharmacokinetic Analyses (Page 22)

Rational: This section has been amended 1) to add details in case of missing items, 2) to drop some visits in analysis per design, and 3) to include the weights for multiple comparison procedure.

This section now reads as follows:

5.1.1.3 Efficacy and Pharmacokinetic Analyses

Efficacy analyses will be based on the ITT Population. Pharmacokinetic endpoints and analyses will not be described in this SAP.

The following efficacy assessments and terms are defined:

Table 5-11 Efficacy Assessments

Assessment/Term	Description
Montgomery-Åsberg Depression Rating Scale (MADRS)	10-Item, clinician-rated scale to evaluate the patient's depressive symptomatology. Each item is scored on a 7-point scale, from 0 to 6 with a score of 0 reflecting no symptoms and a score of 6 reflecting symptoms of maximum severity. The MADRS total score is the sum of the 10 items from the MADRS at each visit if no item has missing score. If the number of missing items at a visit is 2 or less, the MADRS total score will be calculated as (sum of nonmissing items) × (total number of items) / (number of nonmissing items). If more than 2 items of the MADRS are missing, then the total score will be set to missing.
Clinical Global Impressions–Severity (CGI-S)	Clinician-rated scale that measures the overall severity of a patient's illness in comparison with the severity of other patients the physician has observed. The patient is rated on a scale from 1 to 7 with 1 indicating a "normal state" and 7 indicating "among the most extremely ill patients.".
[redacted]	[redacted]

Baseline assessments for applicable efficacy endpoints defined as follows:

5.1.1.3.1 Efficacy Endpoint

The following efficacy variables will be analyzed as described below.

Table 5-13 Efficacy Analyses

Assessment	Parameter	Timing	Methodology
MADRS	Change from baseline in MADRS total score to Week 6	Treatment Period: Week 6	CFB MMRM
CGI-S	Change from baseline in CGI-S score by visit	Treatment Period: Weeks 1, 2, 4, and 6	CFB MMRM

[1] Analysis visits defined in Section 6.2.1.

5.1.1.3.2 Multiple Comparisons Procedure for Primary and Secondary Endpoints

Procedure:

To control the overall type I error rate for multiple comparisons of 2 active doses versus placebo at Week 6 for the primary and secondary efficacy parameters, the parallel gatekeeping procedure will be implemented (Chen et al, 2005).

Hypotheses:

The 4 null hypotheses will be grouped into 2 families denoted by F1 and F2, with F1 consisting of 2 null hypotheses for comparisons of the 2 active doses versus placebo in change from baseline to Week 6 in MADRS total score (i.e., $F_1 = \{H_{11}, H_{12}\}$), and F2 consisting of 2 null hypotheses for comparisons of the 2 active doses versus placebo in change from baseline to Week 6 in CGI-S score (i.e., $F_2 = \{H_{21}, H_{22}\}$). Family F1 will serve as the parallel gatekeeper for F2.

Testing:

The matched gatekeeping procedure utilizes the special logical relationship between the primary and the secondary parameters to enhance the power of statistical testing. The secondary efficacy parameter will be tested at a specific dose only if the corresponding primary efficacy parameter is statistically significant. The Simes test will be used to derive the local p-values for the interaction hypotheses. The adjusted p-values for the 4 elementary hypotheses will be calculated based on the closed testing principle. Statistical significance will be determined by comparing the adjusted p-values to $\alpha = 0.050$.

Weights for Intersection Tests in the Matched Parallel Gatekeeping Procedure

Intersection hypothesis	Weights			
	H_{11}	H_{12}	H_{21}	H_{22}
$H_{11} \cap H_{12} \cap H_{21} \cap H_{22}$	0.5	0.5	0	0
$H_{11} \cap H_{12} \cap H_{21}$	0.5	0.5	0	0
$H_{11} \cap H_{12} \cap H_{22}$	0.5	0.5	0	0
$H_{11} \cap H_{12}$	0.5	0.5	0	0
$H_{11} \cap H_{21} \cap H_{22}$	0.5	0	0	0.5
$H_{11} \cap H_{21}$	1	0	0	0
$H_{11} \cap H_{22}$	0.5	0	0	0.5
H_{11}	1	0	0	0
$H_{12} \cap H_{21} \cap H_{22}$	0	0.5	0.5	0
$H_{12} \cap H_{21}$	0	0.5	0.5	0
$H_{12} \cap H_{22}$	0	1	0	0
H_{12}	0	1	0	0
$H_{21} \cap H_{22}$	0	0	0.5	0.5
H_{21}	0	0	1	0
H_{22}	0	0	0	1

[redacted]

[redacted]

[redacted]

			~~PCS Criteria~~	
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	
	[redacted]	[redacted]	[redacted]	[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted] [redacted]

[redacted]	[redacted]		
	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted] ~~Crystals~~

[redacted]

[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]

[redacted]

[redacted]

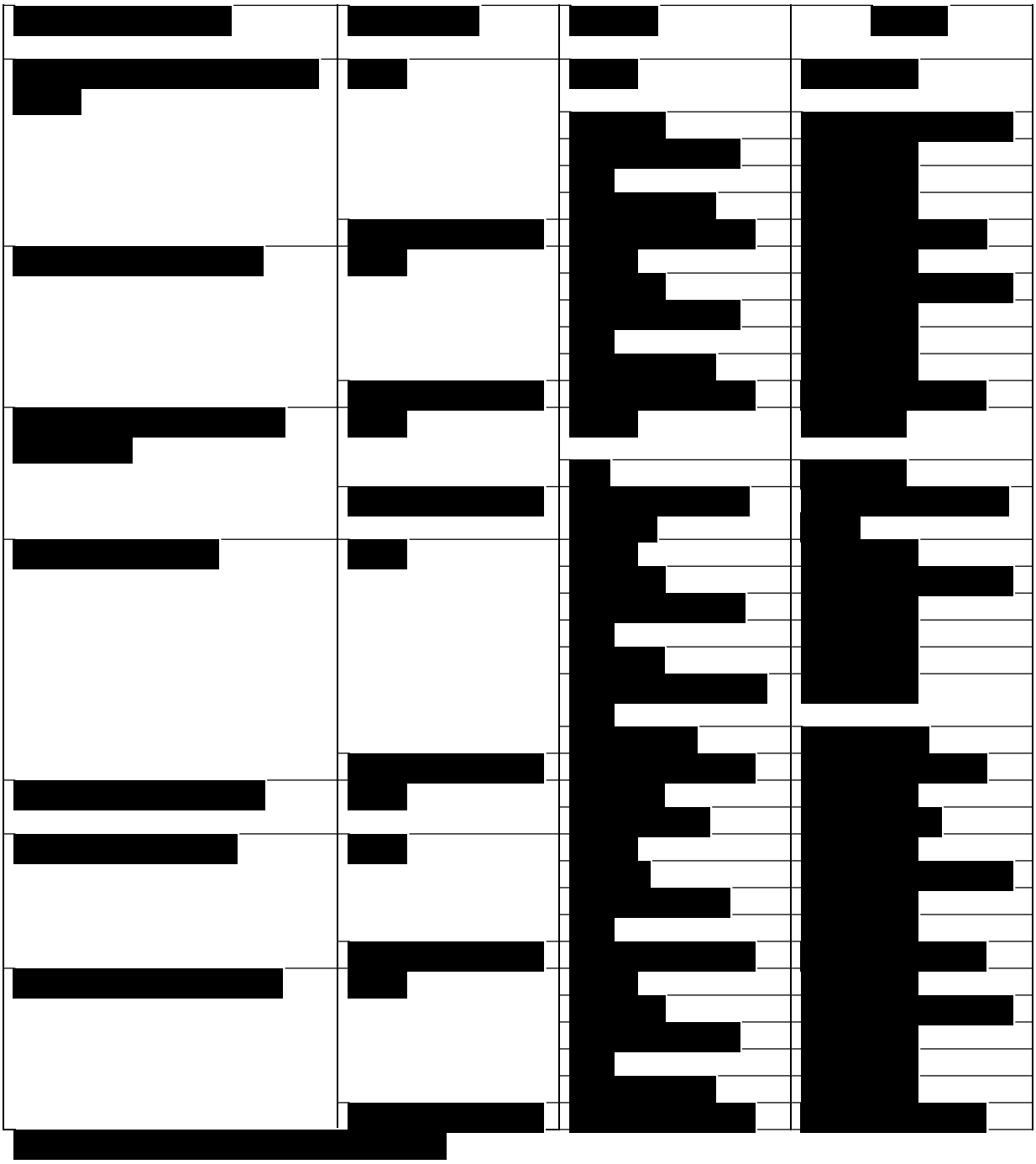
[redacted]

[redacted]

[redacted]

[redacted]

[redacted]



[redacted]

[redacted]

[redacted]

[redacted]

[redacted]	[redacted]	[redacted]
[redacted]		
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]		
[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]
	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]
	[redacted]	[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

		[redacted]	
[redacted]		[redacted]	[redacted]
		[redacted]	[redacted]
		[redacted]	[redacted]

[redacted]