Janssen Research & Development

Statistical Analysis Plan

A Randomized, Double-blind, Placebo-Controlled, Parallel-Group, Multi-Center Study Investigating the Efficacy, Safety, and Tolerability of JNJ-42165279 in Subjects with Major Depressive Disorder with Anxious Distress

Protocol 42165279MDD2001; Phase 2a

JNJ-42165279

Status:	Approved
Date:	14 February 2019
Prepared by:	Janssen Research & Development, a division of Janssen Pharmaceutica NV
Document No.:	EDMS-ERI-180165119

Compliance: The study described in this report was performed according to the principles of Good Clinical Practice (GCP).

TABLE OF CONTENTS

ABBREVIATIONS

AEA	anandamide
AE	adverse event
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
AST	aspartate aminotransferase
ATRQ	Antidepressant Treatment History Questionnaire
BLQ	below the limit of quantification
BMI	body mass index
CGI	clinical global impression
CPFQ	Cognitive and Physical Functioning Questionnaire
CRF	case report form
C-SSRS	Columbia Suicide Severity Rating scale
CUDOS-A	clinically useful depression outcome scale
DBP	diastolic blood pressure
DRC	Data Review Committee
DSM	Diagnostic and Statistical Manual
ECG	electrocardiogram
HAM-A	Hamilton Anxiety Rating Scale
HAM-D	Hamilton Rating Scale for Depression
HDRS	Hamilton Depression Rating Scale
HPA	hypothalamic–pituitary–adrenal
ITT	intention to treat
eITT	enriched intention to treat
fITT	full intention to treat
IVRS	interactive voice response system
IWRS	interactive web response system
LLN	Lower limit of normal
MDD	major depressive disorder
MedDRA	Medical Dictionary for Regulatory Activities
MINI	mini international neuropsychiatric interview
MMRM	mixed model for repeated measures
MOS	Medical Outcomes Study
PEA	palmitoylethanolamide
PK	pharmacokinetic(s)
PSS	Perceived Stress Scale
SAE	serious adverse event
SAP	Statistical Analysis Plan
SATE	Self-Assessment of Treatment Experience
SD	standard deviation
SE	standard error
SHAPS	Snaith–Hamilton Pleasure Scale
SI	Standard international
SIGH-A	Structured Interview Guide For The Hamilton Anxiety Scale
SNRI	serotonergic/noradrenergic reuptake inhibitor
SSRI	selective serotonin reuptake inhibitor
SBP	systolic blood pressure
TEAE	treatment-emergent adverse event
TEMA	Treatment-emergent markedly abnormal
TFL	Tables, figures and listings
ULN	Upper limit of normal
ULQ	under the limit of quantification

1. INTRODUCTION

This statistical analysis plan (SAP) contains definitions of analysis sets, derived variables, and statistical methods for final analyses of efficacy, biomarker and safety data.

1.1. Trial Objectives

Primary Objectives:

The primary objective of this study is to evaluate the efficacy in terms of reduction of symptoms of depression and anxiety, as assessed by the change from baseline on a 17-item Hamilton Depression Rating Scale ($HDRS_{17}$), and overall safety and tolerability of treatment with adjunctive JNJ-42165279 compared to placebo in subjects with major depressive disorder (MDD) with anxiety symptoms who have had inadequate response to treatment with a selective serotonin reuptake inhibitor (SSRI) or serotonergic/noradrenergic reuptake inhibitor (SNRI).

Secondary Objectives:

- To assess the efficacy of JNJ-42165279 on core symptoms of anxiety ($HAM\text{-}A_6$).
- To assess the efficacy of JNJ-42165279 using dimensional analyses of both anxiety and depression ($HDRS_{17}$ and SIGH-A).
- To assess the efficacy of JNJ-42165279 on the response and remission of depressive and anxiety symptoms (derived from $HDRS_{17}$ and SIGH-A).
- To assess the plasma pharmacokinetic (PK) profile of JNJ-42165279 administered as once daily (qd) in male and female subjects with MDD with anxious distress using a population PK approach and explore its relationship with efficacy and safety parameters.

Exploratory Objective:

- To evaluate the impact of treatment with adjunctive JNJ-42165279 compared to placebo on Clinical Global Impression (CGI) and exploratory clinician-rated and patient-reported assessments (CUDOS-A, CPFQ, SHAPS, MOS Sleep-R, PSS and Self-Assessment of Treatment Experience).
- To explore for CCI [redacted] in subjects with CCI [redacted] features and subjects with CCI [redacted], as reported on the CCI [redacted] CCI [redacted].
- To explore CCI [redacted] (including but not limited to CCI [redacted] and CCI [redacted] that may be related to CCI [redacted] of JNJ-42165279.

- To explore the relationship between plasma PK and plasma concentrations of FAAs (anandamide [AEA], palmitoylethanolamide [PEA] and oleoylethanolamide [OEA]) in subjects with MDD with anxious distress.

1.2. Trial Design

This is a multi-center, double-blind, placebo-controlled, randomized, parallel-group study in subjects with MDD with anxious distress who have had inadequate response to SSRI/SNRI treatment.

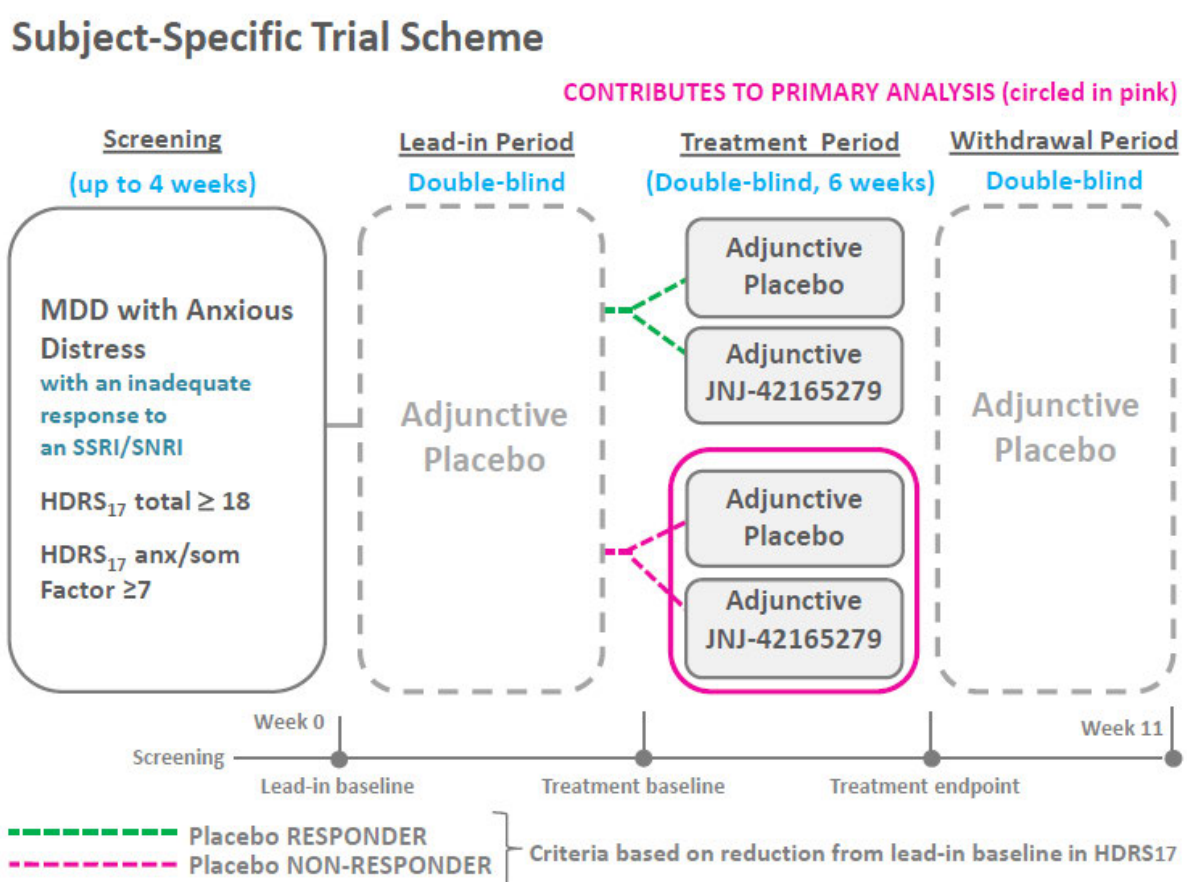
Subjects who meet the inclusion and exclusion criteria and are enrolled will be maintained on the SSRI/SNRI treatment throughout the study to determine whether adjunctive treatment with JNJ-42165279 can reduce symptoms of MDD with anxious distress.

For each subject, the study will consist of three phases: a screening phase of up to 4 weeks, a treatment phase lasting up to 11 weeks, and a 3-week post-treatment (follow up) phase. The treatment phase of the trial will consist of 3 periods. The first period is a placebo lead-in of 3 weeks for all subjects, after which subjects will enter the double-blind treatment period when they will be randomly assigned to JNJ-42165279 or continuation on placebo for 6 weeks. Subjects who successfully complete the double-blind treatment period will be treated with placebo during a 2-week withdrawal period. The total study duration for each subject will be approximately 18 weeks.

The end of study is defined as the date of the last visit of the last subject undergoing the trial. Investigators and subjects will remain blinded to the exact duration of the subject lead-in and withdrawal period during the entire study duration.

Approximately 143 subjects will enter the double-blind placebo lead-in period. At the end of the lead-in period, response status of the subjects will be assessed according to the double-blind response criteria based on reduction in $HDRS_{17}$ relative to lead-in baseline. Both lead-in placebo responders and lead-in placebo non-responders will be randomly assigned in a 1:1 ratio to either 25 mg of JNJ-42165279 or placebo in the treatment period. The primary efficacy analyses will be based on the treatment period data from the lead-in non-responders (94 subjects).

An overall flow diagram of the study is shown in the figure below.



1.3. Statistical Hypotheses for Trial Objectives

The primary efficacy endpoint is the change from baseline in the $HDRS_{17}$ total score at Week 6. The null hypothesis is that there is no difference between treatment with adjunctive JNJ-42165279 and placebo in the improvement of symptoms of depression and anxiety, in subjects with MDD with anxiety symptoms who have had an inadequate response to an SSRI/SNRI, based on the primary efficacy endpoint.

1.4. Sample Size Justification

The sample size for the study is determined based on the assumption of a minimally relevant treatment effect size of 0.45 in the mean change from baseline to the endpoint in $HDRS_{17}$ total score between JNJ-42165279 treatment group and placebo. The assumed effect size is based on review of the literature looking at the treatments for MDD and MDD with anxious distress (Lee 2004 and McIntyre 2007) taking into account that the effect size in enriched population is expected to be higher. A standard deviation of 7.5 in the change in $HDRS_{17}$ total score from baseline is estimated based on previously conducted clinical trial in similar patient population (40411813DAX2001) allowing for a higher variability.

The effect size of 0.45 (under SD assumption of 7.5) translates to a treatment difference between JNJ-42165279 and placebo at endpoint of 3.5 points. Detection of this magnitude of effect size with a power of 90% at an overall 1-sided significance level of 0.20, requires 45 subjects in each treatment group.

Justification of alpha and beta

The choice of alpha and beta (1-power) for this Phase 2 study was made in order to increase sensitivity for detecting a therapeutic signal while also maintaining a modest sample size. Thus, power was set to a high value (power 90%; beta 10%) but the type 1 error rate was specified at 1-sided alpha 0.20. This choice is supported by a recent publication by Lindborg et al in which authors note that type 1 and 2 error levels commonly employed in Phase 3 study designs (simple hypothesis tests with 2-sided alpha 0.05 and beta 0.2) are suboptimal for Phase 2, and that switching these values (increasing alpha while decreasing beta) can increase Phase 2 productivity and reduce the risk of rejecting a compound with significant therapeutic potential.

Adjusting for dropouts

When adjusted for a drop-out rate of approximately 3% of subjects who will have no post treatment baseline efficacy measurement, the required number of subjects to be randomized in the treatment period is 94. To achieve this, the estimated number of subjects to enter the lead in period is 140, after adjusting for an estimated placebo response rate of 25% and dropout rate of 10% during the lead-in period.

To replace 3 subjects who prematurely stopped the study when the study was put on hold, the total number of subjects entering the study will be increased from approximately 140 to 143.

Unblinded Data Review

In January 2016, Janssen decided to suspend the 42165279MDD2001 study because of severe safety issues that occurred for several healthy volunteers in a phase 1 study sponsored by Bial, in which a FAAH inhibitor had been administered. The FDA placed a clinical hold on all trials being conducted with a FAAH inhibitor. The 42165279MDD2001 study was unblinded and the randomization codes were provided to the investigators so that the subjects could be informed about their treatment. Two subjects had been randomized into the double-blind treatment; one to placebo and one to JNJ-42165279, and one subject was still in the lead-in period when the study was suspended. Safety information was listed for these 3 subjects. Following an investigation of the Bial compound, it was determined that the serious adverse events were not related to FAAH inhibition. The FDA lifted the hold on 22 September 2016 and Janssen senior management approved resuming study 42165279MDD2001 in December 2016. The protocol was amended to replace the 3 subjects who prematurely stopped the study when it was put on hold.

Interim Analysis

No interim analysis is foreseen. However, depending on the recruitment, an unblinded review of the data might be conducted for subjects who completed the treatment periods by an internal Data Review Committee (DRC). The DRC may decide to terminate the study after review of the safety data. If any such review is conducted it will be documented in an Early Medicine DRC charter prior to the unblinding. The constitution of the Data review committee will be documented in the Early Medicine Data Review Committee charter and may include sponsor study team members.

Sample Size Re-Estimation

Blinded data review for purpose of sample size re-estimation may be performed after 75% of the subjects are randomized. Sample size may be re-adjusted if observed SD substantially deviates from the hypothesized or if the lead-in response and dropout rate substantially deviate from the assumed. Maximal number of subjects to be enrolled in the trial will not surpass 180.

1.5. Randomization and Blinding

1.5.1. Procedures for Randomization

Central randomization will be implemented in this study. Subjects will be randomly assigned to one of two treatment groups based on the first of two computer-generated randomization schedules prepared before the study by, or under the supervision of the sponsor. The randomization will be balanced by using randomly permuted blocks stratified by (pooled) center and lead-in response status.

Exact timing of randomization for each subject will be blinded. The interactive voice/web response system (IVRS/IWRS) will assign a unique treatment code, which will dictate the treatment assignment and matching study drug kits for the subject. The requestor must use his or her own user identification and personal identification number each time when contacting the IVRS/IWRS and will then give the relevant subject details to uniquely identify the subject.

1.5.2. Blinding

Blinded treatment will be used to reduce potential bias during data collection and evaluation of clinical endpoints. The investigator will not be provided with randomization codes. The codes will be maintained within the IVRS/IWRS, which has the functionality to allow the investigator to break the blind for an individual subject.

Data that may potentially unblind the treatment assignment (e.g., study medication plasma concentrations, plasma biomarkers) will be handled with special care to ensure that the integrity of the blind is maintained and the potential for bias is minimized. This can include making special provisions, such as segregating the data in question from view by the investigators, clinical team, or others as appropriate until the time of database lock and unblinding.

In general, randomization codes will be disclosed fully only if the study is completed and the clinical database is closed. However, if an interim data review by the DRC is specified, the randomization codes and, if required, the translation of randomization codes into treatment and control groups will be disclosed to those authorized and only for those subjects included in the interim review.

2. GENERAL ANALYSIS DEFINITIONS

The statistical analysis will be performed using SAS® version 9.3 or higher.

The analysis will include descriptive statistics, frequency tabulations, statistical analyses, graphical presentations and data listings.

Unless otherwise specified, descriptive statistics of continuous parameters will be summarized using the following statistics: number of observations, mean, standard deviation (SD), standard error (SE), minimum, median and maximum. Frequency tabulations will display the number of subjects and corresponding percentage, where the incidence is based on the number of subjects in each treatment group.

Treatment Group Names

The treatment group names will appear in the tables, figures and listings (TFLs) as:

- Placebo
- JNJ-42165279 25mg

Lead-In Placebo Responders

Lead-in placebo responders are defined as subjects with ≥40% improvement from lead-in baseline in the $HDRS_{17}$ total score at the end of the lead-in period.

Data from eCRF will be used to determine if subjects are lead-in placebo responders, since this reflects the final, clean data. For some subjects, incorrect $HDRS_{17}$ total scores were fed into the IWRS system, and these scores were updated afterwards in the eCRF system (after randomization had occurred, and source data verification revealed that data had been wrongly entered into the IWRS system). A listing will be created with subjects who have $HDRS_{17}$ total score discrepancies between the IWRS system and the eCRF system at lead-in baseline and at the end of the lead-in period. Data from the eCRF system will be used for the analysis.

Initially, at the time of protocol development, the lead-in placebo response definition was set at ≥50% improvement. However, when subject enrollment had started, it became apparent that the expected number of lead-in placebo responders was lower than what was anticipated (0 out of 14 subjects, while it was predicted to have 2-3 subjects). It was clear that out of the first set of 14 subjects, two subjects showed substantial improvement (40% and 41%) but they did not meet the ≥50% improvement threshold. In order to make sure that the enrichment strategy would work, it was decided to update the definition of lead-in placebo response from ≥50% improvement to ≥40% improvement. In the analysis, we will apply the 40% improvement definition to all subjects, and a listing will be created to identify the two subjects who were randomized using the 50% improvement definition. Section 3.2 describes how lead-in placebo response will be summarized.

2.1. Pooling Algorithm for Analysis Centers

Centers will be pooled by country for the statistical analysis.

2.2. Analysis Sets

2.2.1. All Subjects Analysis Set

The **All Subjects Analysis Set** will include all subjects who were screened.

This analysis set will be used for all listings and for calculating and listing the study execution period.

2.2.2. All Enrolled Analysis Set

The **All Enrolled Analysis Set** will include all subjects who were enrolled into the Lead-in period, regardless of whether treatment was received.

This analysis set will be used for summarizing study disposition.

2.2.3. Safety Analysis Set

The **Safety Analysis Set** will be defined for both enriched and full study populations.

- The Safety Analysis Set is defined as all enrolled subjects who received at least one dose of study medication in the Lead-in Period.

- The Enriched Safety Analysis Set is defined as all enrolled subjects who received at least one dose of study medication in the Lead-in Period and who were lead-in placebo non-responders.

The safety analysis set will be used for summarizing safety information. In addition, the enriched safety analysis set will be used for summarizing adverse events.

2.2.4. Efficacy Analysis Sets

The **ITT Analysis Set** will be defined for both enriched and full study populations. At the time of study suspension, two subjects who were randomized had to withdraw early from the study (1 placebo subject, 1 JNJ-42165279 subject, both placebo non-responders). These two subjects will not be included in the ITT analysis sets.

- The Enriched ITT (eITT) Analysis Set is defined as all enrolled subjects who were randomized into the double-blind treatment period, who were lead-in placebo non-responder, who received at least one dose of double-blind study medication and have at least one post-baseline $HDRS_{17}$ assessment during the double-blind treatment period.
- The Full ITT (fITT) Analysis Set is defined as all enrolled subjects who were randomized into the double-blind treatment period, who received at least one dose of double-blind study medication and have at least one post-baseline $HDRS_{17}$ assessment during the double-blind treatment period.

Efficacy analyses will be summarized for both the eITT and fITT analysis sets.

2.2.5. Per Protocol Analysis Set

The **Per Protocol Analysis Set** will include all eITT subjects who did not have a major protocol deviation during the study. A sensitivity analysis of the primary efficacy parameter will be performed for the per protocol analysis set.

2.2.6. Biomarker Analysis Set

The **Biomarker Analysis Set** will include all enrolled subjects who were randomized in the double-blind treatment period, who received at least one dose of double-blind study medication and have biomarker data available in the double-blind treatment period. The biomarker analysis set will be used for summarizing biomarker results.

2.3. Study Day and Relative Day

Screening Period

Study day in the Screening Period is calculated relative to the Lead-in Period start date. All study days in the Screening period will be negative, there is no Study Day 0:

Study Day measurement date – Lead-in Period start date

Lead-in, Double-Blind Treatment and Withdrawal Periods

Study day is calculated relative to the treatment start date in the corresponding study phase. By definition, there is no Study Day 0.

Study Day measurement date – treatment start date + 1

Follow-Up Period

Study day is calculated relative to the Follow-Up Period start date. By definition, there is no Study Day 0.

Study Day measurement date – follow-up period start date + 1

2.4. Analysis Timepoints and Phases

Please refer to the time and events schedule in the latest version of the protocol for analysis timepoints per assessment.

Baseline (reference time point): for each of the Lead-in, DB Treatment and Withdrawal Period, the reference assessment, i.e. baseline, is the measurement the closest, and prior to the first study medication administration of that period. The baseline measurement may be taken from a scheduled or unscheduled visit.

End Point for double-blind treatment period: end point is defined as the last non-missing post-baseline observation during the double-blind treatment period. End point will be defined only for the double-blind treatment period.

Analysis phases will be constructed as follows:

Analysis Phase	Start Date/time	End Date/time
Screening	00:00 of the date of signing the informed consent	1 minute before the first study medication administration in the Lead-in Period
Lead-in Treatment Period	Date/time of the first study medication administration in the Lead-in Treatment Period	1 minute before the first study medication administration in the Double-Blind Treatment Period

		If the subject does not enter the Double-blind Treatment Period then use 23:59 of the day before the date of the first Follow-Up visit
Double-Blind Treatment Period	Date/time of the first study medication administration in the Double-Blind Treatment Period If the subject takes the morning dose on the day of randomization prior to the randomization (i.e. a dose of the lead-in medication), the time of morning intake of the next day is not known. The start date of the double-blind treatment period will be the day after the day of randomization, and the time of intake will be imputed as 0:00.	1 minute before the first study medication administration in the Withdrawal Treatment Period. If the subject does not enter the Withdrawal Treatment Period then use 23:59 of the day before the start of the Follow-Up period.
Withdrawal Treatment Period	Date/time of the first study medication administration in the Withdrawal Period	Day 77 visit or early withdrawal visit for subjects who dropout early.
Follow-up	00:00 of the date after the Day 77 visit or early withdrawal visit for subjects who dropout early.	23:59 of the day of study termination (date of last contact)

The last phase, whichever it is for a subject, always ends on 23:59 of the day of study termination (last contact).

In case of a dropout during a treatment phase: the dropout subject will have a follow-up phase after this phase only if a follow-up visit (as defined in the protocol) was performed.

The number of days per phase will be defined as: *day of phase end – day of phase start + 1*.

2.5. Visit Windows

As subjects do not always adhere to the protocol visit schedule, the following rules are applied to assign actual visits to analysis visits. Listed in Table 1 are visit windows and the target days for each visit. The reference day is Study Day 1 for each of the lead-in, the double-blind treatment and the withdrawal period.

If a subject has 2 or more actual visits in one visit window, the visit closest to the target day will be used as the protocol visit for that visit window. The other additional visit(s) will not be used in the summaries or analyses. If 2 actual visits are equidistant from the target day within a visit

window, the later visit is used. If a visit window has no scheduled visits but does have unscheduled visits, then the unscheduled visit closest to the scheduled visit will be used.

All assignments will be made in chronological order. Once a visit date is assigned to a visit window, it will no longer be used for a later time point except for the endpoint. Listed below (Table 1) are the visit windows and the target days (if applicable) for each visit defined in the protocol.

Table 1 – Visit Windows

Parameter	Scheduled Visit Number	Time Interval (label on output)	Time Interval (Day)[1]	Target Time Point (Day) by Period	Analysis Phase
$HDRS_{17}$, C-SSRS	1	Screening	< 1	-28 to -1	Screening
	2	Lead-in Baseline	<=1	1	Lead-in
	3	Lead-in Week 1	2 to 10	7	Lead-in
	4	Lead-in Week 2	11 to 17	14	Lead-in
	5	Lead-in Week 3	18 to 24	21	Lead-in
	5	DB Treatment Baseline	<=1	1	DB Treatment
	6	DB Treatment Week 1	2 to 10	7	DB Treatment
	7	DB Treatment Week 2	11 to 21	14	DB Treatment
	8	DB Treatment Week 4	22 to 35	28	DB Treatment
	9	DB Treatment Week 6	36 to 49	42	DB Treatment
		DB Treatment End Point[2]			DB Treatment
	9	Withdrawal Baseline	<=1	1	Withdrawal
	10	Withdrawal Week 2	2 to 20	14	Withdrawal
CPFQ, CUDOS-A, SHAPS, MOS Sleep-R, PSS	2	Lead-in Baseline	<=1	1	Lead-in
	3	Lead-in Week 1	2 to 14	7	Lead-in
	5	Lead-in Week 3	15 to 28	21	Lead-in
	5	DB Treatment Baseline	<=1	1	DB Treatment
	7	DB Treatment Week 2	2 to 21	14	DB Treatment
	8	DB Treatment Week 4	22 to 35	28	DB Treatment
	9	DB Treatment Week 6	36 to 49	42	DB Treatment
		DB Treatment End Point[2]			DB Treatment
	9	Withdrawal Baseline	<=1	1	Withdrawal
	10	Withdrawal Week 2	2 to 20	14	Withdrawal
SATE	10	Withdrawal Week 2	2 to 20	14	Withdrawal
SIGH-A	2	Lead-in Baseline	<=1	1	Lead-in
	3	Lead-in Week 1	2 to 10	7	Lead-in
	4	Lead-in Week 2	11 to 17	14	Lead-in
	5	Lead-in Week 3	18 to 24	21	Lead-in
	5	DB Treatment Baseline	<=1	1	DB Treatment
	6	DB Treatment Week 1	2 to 10	7	DB Treatment
	7	DB Treatment Week 2	11 to 21	14	DB Treatment

Parameter	Scheduled Visit Number	Time Interval (label on output)	Time Interval (Day)[1]	Target Time Point (Day) by Period	Analysis Phase
	8	DB Treatment Week 4	22 to 35	28	DB Treatment
	9	DB Treatment Week 6	36 to 49	42	DB Treatment
		DB Treatment End Point[2]			DB Treatment
	9	Withdrawal Baseline	<=1	1	Withdrawal
	10	Withdrawal Week 2	2 to 20	14	Withdrawal
CGI-I	3	Lead-in Week 1	2 to 10	7	Lead-in
	4	Lead-in Week 2	11 to 17	14	Lead-in
	5	Lead-in Week 3	18 to 24	21	Lead-in
	5	DB Treatment Baseline	<=1	1	DB Treatment
	6	DB Treatment Week 1	2 to 10	7	DB Treatment
	7	DB Treatment Week 2	11 to 21	14	DB Treatment
	8	DB Treatment Week 4	22 to 35	28	DB Treatment
	9	DB Treatment Week 6	36 to 49	42	DB Treatment
		DB Treatment End Point[2]			DB Treatment
	9	Withdrawal Baseline	<=1	1	Withdrawal
	10	Withdrawal Week 2	2 to 20	14	Withdrawal
Physical Examination	1	Screening	< 1	-28 to -1	Screening
	10	Withdrawal Week 2	2 to 20	14	Withdrawal
	11	Follow Up	> DB or >Withdrawal	7 to 21	Follow Up
Neurological Examination	1	Screening	< 1	-28 to -1	Screening
	7	DB Treatment Week 2	2 to 21	14	DB Treatment
	9	DB Treatment Week 6	22 to 49	42	DB Treatment
	10	Withdrawal Week 2	2 to 20	14	Withdrawal
Vital Signs, Body Temperature, Body Weight, Clinical Laboratory Assessments, 12-lead ECG	1	Screening	<1	-28 to -1	Screening
	2	Lead-in Baseline	<=1	1	Lead-in
	3	Lead-in Week 1	2 to 14	7	Lead-in
	5	Lead-in Week 3	15 to 28	21	Lead-in
	5	DB Treatment Baseline	<=1	1	DB Treatment
	7	DB Treatment Week 2	2 to 21	14	DB Treatment
	8	DB Treatment Week 4	22 to 35	28	DB Treatment
	9	DB Treatment Week 6	36 to 49	42	DB Treatment

Parameter	Scheduled Visit Number	Time Interval (label on output)	Time Interval (Day)[1]	Target Time Point (Day) by Period	Analysis Phase
		DB Treatment End Point[2]			DB Treatment
	9	Withdrawal Baseline	<=1	1	Withdrawal
	10	Withdrawal Week 2	2 to 20	14	Withdrawal
	11	Follow Up[3]	> DB or >Withdrawal	7 to 21	Follow Up
CCI					

[1] Relative to Study Day 1 in each of the Lead-in, the double-blind treatment and the withdrawal period
[2] End Point is defined as the last non-missing post-baseline measurement during the double-blind treatment period.
[3] For clinical laboratory assessments and ECG, the follow up visit is performed only in case of any clinical significant abnormalities observed at Week 11.

3. SUBJECT INFORMATION

Subject information will be summarized for the Safety Analysis Set and listings will be created for the All Subjects Set.

3.1. Demographics and Baseline Characteristics

The following demographic and baseline characteristics variables will be listed and summarized by treatment group and overall:

- age at screening (years)
- sex
- race and ethnicity
- country
- lead-in response status
- height at baseline (cm)
- weight at baseline (kg)
- BMI at baseline (kg/m²) baseline weight (kg) / (baseline height (m))²
- duration of prior antidepressant treatment (in days)

Baseline disease characteristics will be summarized at lead-in baseline and at double-blind treatment baseline for the following assessments: $HDRS_{17}$ total score, SIGH-A total score, HAM-A6 subscale score, HAM-D6 subscale score and $HDRS_{17}$ anxiety/somatization factor score.

The summary of demographic and baseline characteristics and baseline disease characteristics will be presented for the Safety analysis set, the Enriched Safety analysis set, the Enriched ITT set and the Full ITT set. In addition, the summary for the Safety analysis set will be presented by country and by enrollment prior to/after study suspension.

The following screening and/or pre-dosing test results will be listed:

- urine drug screen abnormalities
- positive serology results
- serum FSH results that are ≤40 IU/L (females only)
- positive alcohol results

If no abnormalities or positive results, an empty listing will be produced stating "No data to report". If a subject has an abnormal finding, all related results will be listed (e.g. all serology results, or all drug screening results).

All medical history data will be presented in a listed.

Results from the Mini-international neuropsychiatric interview (MINI) will be listed.

3.2. Disposition Information

The number of subjects in each analysis set will be presented, and the number of subjects by country and center will be presented for the Safety analysis set and for the Enriched ITT set.

Study completion/withdrawal information will be summarized by treatment group and overall and will be listed. The distribution of study termination reasons will also be summarized.

The distribution of time to study discontinuation will be presented graphically using Kaplan-Meier estimates. The time to discontinuation will be calculated as the duration (in days) from Study Day 1 of the lead-in period (defined in Section 2.4) to study discontinuation. Subjects who complete the study will be censored at the study disposition date.

Study completion and withdrawal information will be presented for the Safety analysis set and the Enriched ITT set and the information for each subject will also be presented in a listing.

The number and percentage of subjects who enter lead-in period, the double-blind treatment period and the withdrawal period and who complete the lead-in period, the double-blind treatment period and the withdrawal period will be summarized.

The number and percentage of lead-in placebo responders will be summarized overall, by country and by site. A cumulative distribution plot of the percent reduction in $HDRS_{17}$ total score from lead-in baseline at the end of the lead-in period will be provided overall and by country. To explore the duration of response for lead-in placebo responders, a Kaplan-Meier plot will be presented for the distribution of time to worsening (i.e. no longer meeting ≥40% improvement) after the lead-in period. Time to worsening will be calculated as the duration (in days) from Study Day 1 of the double-blind treatment period (defined in Section 2.4) to the first visit where <40% improvement in $HDRS_{17}$ total score is observed. Subjects who do not reach the worsening criterion will be censored at the study disposition date.

3.3. Treatment Compliance

The number of tablets dispensed and returned at each visit will be listed.

3.4. Extent of Exposure

Study drug administration information will be presented for each subject in a data listing.

Treatment duration is defined as the number of days between the first and last day of study drug in the treatment phase (including lead-in, double-blind treatment and withdrawal): last day of study drug administration – first day of study drug administration + 1.

Descriptive statistics of treatment duration will be presented. A frequency distribution for the treatment period only will also be provided using the following categories: <1 week (<7 days), 1-<2 weeks (7-<14 days), 2-<3 weeks (14-<21 days), 3-<5 weeks (21-<35 days), 5-<7 weeks (35-<49 days), 7-<9 weeks (49-<63 days), 9-<11 weeks (63-<77 days) and ≥11 weeks (≥77 days).

3.5. Protocol Deviations

All major protocol deviations will be listed.

3.6. Prior and Concomitant Medications

Prior therapies and concomitant medications (relative to the double-blind treatment period) will be tabulated in 3 separate summaries:

1. **Prior medication** – any medication which starts and ends before the date/time of first study drug administration in the Double-blind Treatment Period
2. **Prior and concomitant medication** – any medication which starts before the date/time of first study drug administration in the Double-blind Treatment Period and ends after the date/time of first study drug administration in the Double-blind Treatment Period
3. **Concomitant medication** – any medication which starts after the date/time of first study drug administration in the Double-blind Treatment Period

If a prior or concomitant therapy record is missing components of its start and/or stop dates (day and/or month and/or year) then missing data will not be imputed, however, the following approaches will apply:

- If the start year is the same as the year of first drug administration but the start day and month is missing, the medication will be considered as having started before the study.
- If the start month and year is the same as the month and year of first drug administration but the start day is missing, the medication will be considered as having started before the study.
- In case of a completely missing start date, the medication will be considered as having started before the study.
- If the end year is the same as the year of last drug administration but the end day and month is missing, the medication will be considered as ongoing at the end of the study.
- If the end month and year is the same as the month and year of first drug administration but the end day is missing, the medication will be considered as ongoing at the end of the study.
- In case of a completely missing end date, the medication will be considered as ongoing at the end of the study.

In addition, the following will also be listed:

- preplanned surgery/procedures
- antidepressant Treatment History Questionnaire (ATRQ) – any medication that is recorded on the ATRQ form

4. EFFICACY

Efficacy data will be summarized for the ITT Analysis Sets (full and enriched) and listings will be created for the All Subjects Set.

The efficacy variables in this study are listed in Table 2.

Table 2: Efficacy Variables

Endpoint	Scale*
Primary	$HDRS_{17}$ total score
Secondary	$HDRS_{17}$ ≥30% improvement from double-blind treatment baseline on total score $HDRS_{17}$ ≥50% improvement from double-blind treatment baseline on total score (clinical response) $HDRS_{17}$ clinical remission $HDRS_{17}$ anxiety/somatization factor score HAM-D_6 score SIGH-A total score SIGH-A ≥30% improvement from double-blind treatment baseline on total score SIGH-A ≥50% improvement from double-blind treatment baseline on total score HAM-A_6 score CGI-I
Patient Reported Outcomes	Clinically Useful Depression Outcome Scale (CUDOS-A) Cognitive and Physical Functioning Questionnaire (CPFQ) Snaith-Hamilton Pleasure Scale (SHAPS) Medical Outcomes Study Sleep-Revised (MOS Sleep-R) Perceived Stress Scale (PSS) Self-Assessment of Treatment Experience (SATE)

* $HDRS_{17}$ Hamilton Depression Rating Scale; HAM D_6 Hamilton Depression Rating Scale; SIGH A Structured Interview Guide for the Hamilton Anxiety Scale; HAM A_6 Hamilton Anxiety Rating Scale ; CGI I Clinical Global Impression – Improvement.

4.1. Analysis Specifications

4.1.1. Level of Significance

The overall type I error rate for testing the JNJ-42165279 group versus the placebo group for the primary efficacy analysis will be controlled at the 1-sided significance level of 0.20.

4.1.2. Imputation Methods for Missing Items

Imputation of missing individual items will apply only to the $HDRS_{17}$, as described in Section 4.2.1. For all other scales where multiple items are summed to create a total score, if any item of the scale is missing on a visit, the total score for that scale at that visit will be left blank.

4.2. Primary Efficacy Endpoint – $HDRS_{17}$ Total Score

The primary efficacy endpoint is the change from Baseline in $HDRS_{17}$ total score at Week 6 of the double-blind treatment period.

4.2.1. Definition

The $HDRS_{17}$ is a clinician-administered rating scale designed to assess the severity of symptoms in subjects diagnosed with depression (Hamilton M 1960). It is the most widely used symptom severity measure for depression. Each of the 17 items is rated by the clinician on either a 3- or a 5-point scale. Higher scores indicate greater severity of depression. The $HDRS_{17}$ total score will be calculated as the sum of the 17 item scores, and ranges from 0 to 52. If one $HDRS_{17}$ Total Score item is missing, it will be imputed with the closest integer to the average of the remaining items at that time point. If more than 1 item is missing, no imputation will be performed and the $HDRS_{17}$ Total Score will be missing. Imputation of item scores is performed prior to determining the endpoint for the $HDRS_{17}$ Total Score.

4.2.2. Estimand

The primary estimand, the main clinical quantity of interest to be estimated in the study, is defined by the following 4 components:

- **Population:** Subjects with a primary DSM-5 diagnosis of MDD with Anxious Distress
- **Variable:** Change from baseline to Week 6 of the double-blind treatment period in $HDRS_{17}$ total score
- **Intercurrent event:** No intercurrent events to be taken into account
- **Population-level summary:** Difference in mean change from baseline to Week 6 of the double-blind treatment period in $HDRS_{17}$ total score between treatment conditions.

4.2.3. Analysis Methods

4.2.3.1. Primary Analysis

The primary efficacy analysis will be based on the eITT set.

Descriptive statistics of actual values and mean changes from baseline for the $HDRS_{17}$ total score at each scheduled time point will be provided by treatment group for the lead-in period, the double-blind treatment period and the withdrawal period for the eITT set. A mean-SE plot over time will be presented for the $HDRS_{17}$ total score observed values during the lead-in period, the double-blind treatment period and the withdrawal period by treatment group. Listings will be created with $HDRS_{17}$ item scores and the $HDRS_{17}$ total score.

The JNJ-42165279 treatment group will be compared with placebo using the change from double-blind treatment baseline in the $HDRS_{17}$ total score during the double-blind treatment period. The comparison will be performed with a mixed-effects model using repeated measures (MMRM), with time, treatment (placebo, JNJ-42165279) and time-by-treatment interaction as factors, double-blind treatment baseline total $HDRS_{17}$ score as a continuous covariate, and country as categorical covariates. An unstructured variance-covariance matrix will be used. In case of convergence problems, simpler variance-covariance structures such as Toeplitz or AR(1) will be considered. The selection of any of these structures will be determined after exploration of the

observed correlation structure. The treatment-placebo differences will be obtained using the appropriate contrast in the MMRM models at the 6-week endpoint. A one-sided p-value with level of significance of 20% will be derived from the SAS results based on the above mixed-effects model (two-sided p-values at 40% significance level). The reduction in $HDRS_{17}$ total score is expected to be larger after dosing with JNJ42165279, so the one-sided p-value will be calculated as described in the table below.

Calculation of one-sided p-value

Parameter	H_a	Estimated difference Mean $_{JNJ}$ – Mean $_{Pla}$	One-sided p-value
Change in $HDRS_{17}$ total score at Week 6	Mean Chg $_{JNJ}$- Mean Chg $_{Pla}$< 0 (larger reduction expected in JNJ vs pla)	≤ 0	p-value $_{(2\text{-sided test})}$ / 2
		> 0	1 – (p-value $_{(2\text{-sided test})}$ / 2)

Least-square mean changes from baseline (+/-SE) over time during the double-blind treatment period will be presented in a figure.

4.2.3.2. Additional Analysis

To assess the sensitivity of the results of the MMRM analysis of the primary endpoint, an analysis of covariance (ANCOVA) model for the change from baseline to the 6-week double-blind treatment endpoint in $HDRS_{17}$ total score will be carried out. The ANCOVA model will include factors for treatment and (pooled) center and double-blind treatment baseline $HDRS_{17}$ total sore as a covariate. In addition, the same ANCOVA model will also be performed on observed case data at week 6 of the double-blind treatment period.

As a secondary analysis, the MMRM model described in Section 4.2.3.1 will be fitted with age, sex and duration of antidepressant treatment (in days) as additional covariates.

The primary analysis specified in Section 4.2.3.1. will also be performed for the fITT set. Treatment effect for the fITT set will be estimated in 2 manners: using the same MMRM model as for the primary analysis and using the MMRM model with lead-in response status as additional categorical covariate.

To explore the results for the subgroup of placebo responders, descriptive statistics of actual values and mean changes from baseline for the $HDRS_{17}$ total score at each scheduled time point will be provided by treatment group for the lead-in period, the double-blind treatment period and the withdrawal period for the fITT subset of placebo responders. A mean-SE plot over time will also be presented for this subgroup.

4.3. Secondary Efficacy Endpoints

4.3.1. Definition

The following secondary efficacy variables are used to measure depressive symptoms in subjects:

- **$HDRS_{17}$ ≥30% improvement from double-blind treatment baseline on total score**: the percentage change from double-blind treatment baseline is calculated as

$$100 * \frac{(HDRS17\ total\ score\ at\ any\ time\ point - HDRS17\ total\ score\ at\ baseline)}{HDRS17\ total\ score\ at\ baseline}$$

  A ≥30% improvement from double-blind treatment baseline on total score means that the percentage from baseline is ≤ -30 in this calculation.

- **$HDRS_{17}$ ≥50% improvement from double-blind treatment baseline on total score (clinical response)**: the percentage change from double-blind treatment baseline is calculated in the same way as for the ≥30% definition. A ≥50% improvement from double-blind treatment baseline on total score means that the percentage from baseline is ≤ -50 in this calculation.

- **$HDRS_{17}$ clinical remission:** Clinical remission is defined as having a $HDRS_{17}$ total score ≤ 7.

- **$HDRS_{17}$ anxiety/somatization factor score:** The HDRS17 anxiety/somatization factor derived from Cleary and Guy's factor analysis of the HDRS17 scale, includes six items from the original 17-item version: the items for psychic anxiety, somatic anxiety, gastrointestinal somatic symptoms, general somatic symptoms, hypochondriasis, and insight. The HDRS17 anxiety/somatization factor score will be calculated as the sum of the 6 item scores, and ranges from 0 to 18, with higher scores indicating greater severity of symptoms.

- **HAM-D6 subscale score:** A 6-item subscale from the HDRS17 (HAM-D6) will be analyzed as it has been shown to be a uni-dimensional scale that provides information to core depressive symptoms and is sensitive to treatment response (Bech 1975). The six items are: depressed mood, guilt feelings, work and interests, psychomotor retardation, psychic anxiety, and general somatics (tiredness and pains). The HAM-D6 score will be calculated by summing the 6 items scores, and ranges from 0 to 22. Higher scores indicate greater severity of core symptoms.

- **Structured Interview Guide for the Hamilton Anxiety scale (SIGH-A):** The original HAM-A scale assesses the severity of different anxiety-related symptoms (Hamilton 1959; Hamilton 1969). It is the most widely used symptom severity measure for anxiety. As the original HAM-A scale lacks instructions for administration and clear anchor points for the assignment of severity ratings, the structured interview guide version (SIGH-A) is used in the current study. The SIGH-A has been shown to have high inter-rater and test-retest reliability and produced similar but consistently higher (+ 4.2) scores compared to the original HAM-A. The SIGH-A scale consists of 14 items with a score of 0 to 4. Higher scores indicate higher severity (0-absent, 1-mild, 2-moderate, 3-severe, 4-incapacitating). The SIGH-A total score will be calculated by summing the 14 item scores, and ranges from 0 to 56. Higher scores indicate worse results.

- **SIGH-A ≥30% improvement from double-blind treatment baseline on total score**: the percentage change from double-blind treatment baseline is calculated as

$$100 * \frac{(SIGH - A\ total\ score\ at\ any\ time\ point - SIGH - A\ total\ score\ at\ baseline)}{SIGH - A\ total\ score\ at\ baseline}$$

  A ≥30% improvement from double-blind treatment baseline on total score means that the percentage from baseline is ≤ -30 in this calculation.

- **SIGH-A ≥50% improvement from double-blind treatment baseline on total score**: the percentage change from double-blind treatment baseline is calculated in the same way as for the ≥30% definition. A ≥50% improvement from double-blind treatment baseline on total score means that the percentage from baseline is ≤ -50 in this calculation.

- **HAM-A6 subscale score:** The HAM-A6 is a 6-item subscale derived from the original Hamilton Anxiety scale (HAM-A). It comprises five psychic anxiety symptoms: anxious mood, psychic tension, fears, intellectual disturbances, and anxious behavior observed at the interview, as well as one somatic item, muscular tension. The HAM-A6 score will be calculated by summing the 6 item scores, and ranges from 0 to 24. Higher scores indicate greater severity of symptoms.

- **Clinical Global Impression Improvement (CGI-I):** The CGI-I is a 7-point scale that requires the clinician to assess how much the subject's illness has improved or worsened relative to a baseline state at the beginning of the intervention. The CGI-I is rated as: 1 very much improved; 2 much improved; 3 minimally improved; 4 no change; 5 minimally worse; 6 much worse; 7 very much worse. Responders on the CGI-I are defined as subjects with a score of 1 very much improved or 2 much improved.

4.3.2. Analysis Methods

The secondary efficacy endpoints will be analyzed for both the eITT and fITT analysis sets.

For all continuous secondary efficacy endpoints, descriptive statistics of actual values and mean changes from baseline at each scheduled time point will be provided by treatment group for the lead-in period, the double-blind treatment period and the withdrawal period. A mean-SE plot over time will be presented for observed values during the lead-in period, the double-blind treatment period and the withdrawal period by treatment group.

For each of the continuous secondary efficacy endpoints, the JNJ-42165279 treatment group will be compared to placebo using the same MMRM model as for the $HDRS_{17}$ total score.

To explore the results for the subgroup of placebo responders, descriptive statistics of actual values and mean changes from baseline for the continuous secondary endpoints at each scheduled time point will be provided by treatment group for the lead-in period, the double-blind treatment period and the withdrawal period for the fITT subset of placebo responders. Mean-SE plots over time will also be presented for this subgroup.

Frequency tables and figures for the categorical secondary efficacy endpoints will be provided at each scheduled time point by treatment group for the lead-in period, the double-blind treatment period and the withdrawal period. Chi square test will be used to test the overall differences between the treatment groups at week 6 of the double-blind treatment period. If deemed appropriate, Cochran–Mantel–Haenszel test with (pooled) center as stratification factor will be performed.

A cumulative distribution plot of the percent reduction in $HDRS_{17}$ total score from double-blind treatment baseline at Week 6 of the double-blind treatment period will be provided. A similar plot will be provided for the SIGH-A total score percent reduction.

4.4. Patient Reported Outcomes

4.4.1. Definitions

- **Clinically Useful Depression Outcome Scale (CUDOS-A):** The CUDOS contains 18 items assessing all of the Diagnostic and Statistical Manual of Mental Disorders (5^{th} edition) (DSM-5) inclusion criteria for major depressive disorder and dysthymic disorder as well as psychosocial impairment and quality of life. The 16 symptom items are rated on a 5-point Likert scale in order to keep the scale brief. The CUDOS was designed to be brief (completed in less than 3 minutes), quickly scored (in less than 15 seconds), clinically useful (fully covering the DSM-5 symptoms of major depressive disorder and dysthymic disorder), reliable, valid and sensitive to change. The content of the CUDOS-A subscale was based on the DSM-5 criteria for the anxious distress specifier. A CUDOS-A total score will be calculated as the sum of the 5 item scores, and ranges from 0 to 20, with higher scores indicates higher levels of anxious distress.
- **Cognitive and Physical Functioning Questionnaire (CPFQ):** The CPFQ is a brief self-report scale that provides additional information regarding the impact of adjunctive treatment on aspects of cognitive function including attention, memory and mental acuity. Subjects with MDD are often reported to have difficulties with functioning in this area. A CPFQ total score will be calculated as the sum of the 7 item scores, and ranges from 7 to 42, with higher scores indicating more difficulties with cognitive and physical functioning.
- **Snaith-Hamilton Pleasure Scale (SHAPS):** The SHAPS is a 14-item, self-report instrument developed for the assessment of hedonic capacity. The SHAPS was developed to minimize cultural, gender, and age biases in the evaluation of hedonic capacity. It not only measures hedonic tone, but also its absence, i.e. anhedonia. Anhedonia can be a core symptom of depression. Four major domains are covered in the scale, namely interest/pastimes, social interaction, sensory experience, and food/drink. Each of the SHAPS items has a set of four response categories: Strongly Disagree, Disagree, Agree and Strongly Agree, with either of the Agree responses receiving a score of 0 and either of the Disagree responses receiving a score of 1. A SHAPS total score will be calculated as the sum of the 14 item scores, and ranges from 0 to 14. A higher SHAPS total score indicates higher levels of present state of anhedonia.

- **Medical Outcomes Study Sleep-Revised (MOS Sleep-R):** Symptoms of poor sleep commonly occur in anxiety and mood disorders. The MOS Sleep-R is a subject-completed scale containing 12 items that addresses various dimensions of sleep. Ten items are answered on a 6-point scale, where 1="all of the time" and 6="none of the time". One item on sleep latency is answered on a 5 point Likert scale from 1="0-15 minutes" to 5="more than 60 minutes". One item on the duration of sleep allows the subject to write in the number of hours slept per night. The instrument yields six subscales: sleep disturbance, snoring, shortness of breath or headache, sleep adequacy, sleep somnolence, and sleep quantity. The version to be used in this study has a recall period of the 4 past weeks. Quantity of sleep is scored as the average number of hours slept per night. Other subscales scores are converted to a T-score with a mean of 50, standard deviation (SD) of 10 and range of 0 to 100, where higher scores indicate fewer sleep-related problems.
- **Perceived Stress Scale (PSS):** The PSS is the most widely used psychological instrument for measuring the perception of stress. It is a measure of the degree to which situations in one's life are appraised as stressful. The questions in the PSS ask about feelings and thoughts during the last month. In each case, respondents are asked how often they felt a certain way (0 never, 1 almost never, 2 sometimes, 3 fairly often, 4 very often). A PSS total score is obtained by reversing responses (e.g. 0 4, 1 3, 2 2) to the 4 positively stated items (items 4, 5, 7 and 8) and then summing across all items. The PSS total score ranges from 0 to 40 with higher scores indicating higher severity of symptoms.
- **Self-Assessment of Treatment Experience (SATE):** The SATE questionnaire is a 4-item self-report scale designed to provide additional information regarding the subject's subjective experience while taking the treatment.

4.4.2. Analysis Methods

The patient reported outcome efficacy endpoints will be analyzed for both the eITT and fITT analysis sets.

For all continuous patient reported outcome efficacy endpoints, descriptive statistics of actual values and mean changes from baseline at each scheduled time point will be provided by treatment group for the lead-in period, the double-blind treatment period and the withdrawal period. A mean-SE plot over time will be presented for observed values during the lead-in period, the double-blind treatment period and the withdrawal period by treatment group.

Frequency tables and figures for the categorical patient reported outcome efficacy endpoints will be provided at each scheduled time point by treatment group for the lead-in period, the double-blind treatment period and the withdrawal period.

5. SAFETY

All safety analyses will be based on the Safety analysis set, except for the analysis of adverse events, that will also be performed for the enriched Safety analysis set.

All safety data will be listed for the All Subjects Set.

5.1. Imputation Rules for Missing AE Date/Time of Onset/Resolution

A conservative approach will be used to handle the missing dates for adverse events. The rules for estimating incomplete AE onset dates will be as follows:

(1) The missing day of the month will be estimated as follows: If the month and year are known and double-blind study medication started during that month then the estimated date is the start date of double-blind study medication. If the month and year are known and double-blind study medication started prior to that month then the estimated date is the 1st day of the month. If the month and year are known and double-blind study medication started after the month, then no estimation will be done, and the AE will not be considered as treatment emergent for the double-blind phase.

(2) If both the day and the month are missing: No estimation will be performed. However, these AEs will be considered treatment emergent for the double-blind phase and will be included in the double-blind treatment summaries, except for the calculation of duration of the AE. Attempts will be made to get at least the month for the adverse events.

The rules for estimating incomplete AE onset times will be as follows:

The missing times will be estimated as follows: use time 00:00:00 of the given/imputed date for the start date/time of an AE and use 23:59:59 for the end date/time of an AE.

For incomplete AE resolution dates, the rules are:

(1) The missing day of the month will be estimated as follows: If the month and year are known and the study medication was stopped before, or during that month, the estimated date is the last day of the month or the end of the double-blind phase, whichever is earlier. If the study medication stopped after that month then the estimated date is the last day of the month.

(2) If both the day and the month are missing: the estimated resolution date is the end of the double-blind phase.

5.2. Adverse Events

The verbatim terms used in the CRFs by the investigator to identify adverse events (AEs) will be coded using the current Medical Dictionary for Regulatory Activities (MedDRA) maintained by the sponsor.

Therapeutic reach is the number of days after the last dose intake that a subject is still considered to be potentially affected by study drug. For JNJ-42165279, therapeutic reach is defined as 4 days. AEs during the double-blind treatment period will be summarized using the therapeutic reach. Therapeutic reach will only be applied to the end of double-blind treatment period or to subjects who discontinue during the double-blind treatment period.

A treatment-emergent AE (TEAE) is an event that is new in onset or increased in severity following treatment initiation. Any event with a start date and time after the first dose intake will be considered treatment-emergent. An event that starts prior to and ends after the initiation of study medication will be considered treatment-emergent only if the severity increases after the start of medication. An event that starts in the withdrawal or follow up period but within the therapeutic reach of 4 days of the double-blind treatment period will be considered treatment-emergent during the double-blind treatment period. Adverse events with onset date during the withdrawal/follow-up period outside the therapeutic reach of 4 days will be summarized separately.

- The number (%) of subjects with TEAEs will be summarized by system organ class and preferred term for each treatment group and study period. A summary will be created for the most common TEAEs (≥5% incidence in either treatment group). Both of these summaries will also be provided for the enriched Safety analysis set.

- In addition, TEAEs will be summarized by severity and relationship to study drug for each treatment group and study period by system organ class and preferred term. For the summaries of AEs by severity/relationship to study drug, the observation with the most severe occurrence/closest relationship to study drug will be chosen if there is more than one incident of an adverse event reported during the treatment phase by the subject.

- The following TEAEs are considered to be of special interest in this study: severe headache, change in mental states, change in sensory function, change in motor function or coordination. The number (%) of subjects with TEAEs of special interest will be summarized by system organ class and preferred term for each treatment group and study period. Listings will be created with the results of neurological examinations following the incidence of these AEs.

- Data listings will be generated for all AEs, deaths, serious adverse events (SAEs), severe AEs and discontinuations due to AEs and AEs of special interest. These listings will not be limited to TEAEs but will also include any adverse events with onset before the start of study treatment or after the end of study treatment. If there is no data to present then the listing should be created and should state "No data to present".

5.3. Clinical Laboratory Tests

For each study period, descriptive statistics on actual values and changes from respective period baseline at each scheduled time point for hematology, chemistry, urinalysis and coagulation will be provided for the Safety analysis set. Laboratory summaries will be provided in Standard International (SI) units. All laboratory data will be listed for the All Subjects set.

Cross-tabulations (with classes for below, within, and above the central laboratory range) showing the shift of laboratory values from baseline to each scheduled time point and endpoint will be presented by treatment group and study period.

Clinical laboratory test values will be considered "treatment emergent markedly abnormal" (TEMA) using the criteria defined by the Sponsor (Johnson & Johnson Research & Development, LLC) listed in Table 3. The identification of TEMA laboratory values is based on baseline value (in any period) being missing or within the normal range and the corresponding post baseline value being markedly high/low as per the values in Table 3. If post baseline laboratory results are above the upper limit and the baseline value is below the lower limit, then the post baseline abnormality will also be considered TEMA. The same applies to the post baseline value being below the lower limit with the baseline value being above the upper limit.

The incidence of TEMA will be presented by treatment group and study period at each scheduled time point and endpoint.

Table 3: Laboratory Parameters: Criteria for Markedly Abnormal Values

Laboratory Parameter	Markedly Low	Markedly High
Albumin (g/L)	<20	> 2 x ULN
Alkaline phosphatase (U/L)	N/A	> 1.5 x ULN
Alanine aminotransferase (ALT) (U/L)	N/A	> 3 x ULN
Aspartate aminotransferase (AST) (U/L)	N/A	> 3 x ULN
Bicarbonate [mmol/L]	15.1	34.9
Bilirubin (umol/L)	N/A	> 1.5 x ULN
Urea Nitrogen (mmol/L)	N/A	> 3 x ULN
Calcium (mmol/L)	<1.75	>3
Chloride (mmol/L)	<90	>120
Creatine kinase (U/L)	N/A	> 3 x ULN
Creatinine (umol/L)	N/A	> 1.5 x ULN
Gamma glutamyl transferase (U/L)	N/A	> 3 x ULN
Glucose (mmol/L)	<3.33	> 3 x ULN
Lactate dehydrogenase (U/L)	N/A	> 1.5 x ULN
Phosphate (mmol/L)	<0.484	> 1.5 x ULN
Potassium (mmol/L)	<2.8	>5.8
Protein (g/L)	N/A	> 2 x ULN
Prothrombin International Normalized Ratio	N/A	≥1.3
Sodium (mmol/L)	<125.0	>155.0
Hemoglobin (g/L)	<75	> 1.1 x ULN
Platelets (x10e9/L)	<100	> 3 x ULN
Erythrocytes (x10e12/L)	<3	> 2 x ULN
Leukocytes (x10e9/L)	<2.5	> 2.5 x ULN
Hematocrit [fraction]		
male	0.24	0.55
female	0.28	0.5

Graphical presentations will be created for the following liver function tests: alanine aminotransferase, aspartate aminotransferase, albumin, bilirubin, alkaline phosphatase, gamma glutamyl transferase. Liver function test results standardized to the upper limit of normal (based on central laboratory ranges) will be graphically presented over time for each subject. The incidence of liver function tests above the central laboratory normal limit will be presented by treatment group at each scheduled time point and endpoint.

Graphical presentations will also be created for prothrombin international normalized ratio results.

Clinical laboratory tests that meet the criteria for abnormal and markedly abnormal will be listed by subject. Results on urine drug screen, alcohol screen and pregnancy test will be presented in separate listings.

Rule for BQL/UQL (if applicable)

For the measurement that is BQL/UQL, the value will be imputed based on the lowest/highest limit of quantifiable value for the summary table, e.g. if the original result of measurement is marked as '<0.1', the value will be imputed as 0.1 correspondingly. For the listing, however, the original values as stated in the database will be presented.

5.4. Vital Signs and Physical Examination Findings

Descriptive statistics for absolute values and changes from baseline will be presented at each scheduled time point and endpoint by treatment group and study period for body weight, supine systolic blood pressure (mmHg), supine diastolic blood pressure (mmHg), pulse rate (beats per minute) and oral temperature (^{o}C).

Mean-SE plots over time will be presented for vital sign values by treatment group. In addition, scatter plots of post-baseline versus baseline values will be presented by treatment group. All individual vital signs data will be listed.

Abnormal values are defined in Table 4 and will be summarized using frequency distribution tables over time by treatment group and study period. A listing of subjects meeting any of the abnormality criteria will be provided.

Table 4: Vital Signs Abnormality Ranges (Supine and Standing)

	Low	Normal	High
SBP (mmJHg)	<90	90-140	>140
DBP (mmHg)	<50	50-90	>90
Pulse (bpm)	<45	45-90	>90
Temperature (C)	<35.5	35.5-37.5	>37.5

The incidence of subjects who have a treatment-emergent abnormality (as defined in Table 5 below) at each scheduled time point and endpoint will be presented by treatment group and study period.

Table 5: Treatment-Emergent Abnormality Categories for Vital Signs

Vital Parameter	Post-baseline value outside of normal limit if:	
	Abnormally low	**Abnormally high**
Pulse (bpm)	A decrease from baseline of ≥ 15 to a value < 45	An increase from baseline of ≥ 15 to a value > 90
Systolic BP (mmHg)	A decrease from baseline of ≥ 20 to a value < 90	An increase from baseline of ≥ 20 to a value > 140
Diastolic BP (mmHg)	A decrease from baseline of ≥ 15 to a value < 50	An increase from baseline of ≥ 15 to a value > 90

Physical and neurological examination abnormalities will be listed. If a subject has an abnormality noted at any time, all the observations will be listed.

5.5. Electrocardiogram

Triplicate 12-lead ECG will be obtained at Day 1 (predose), otherwise a single, 12-lead ECG will be obtained at each scheduled time point. The ECG variables that will be analyzed are heart rate (HR), ECG intervals (RR, PR, QRS, and QT) as well as corrected QT intervals according to Bazett's formula (QTcB) and Fridericia's formula (QTcF).

Summary tables for actual values and changes from baseline will be presented by treatment group and period at each scheduled time point and at endpoint. Mean-SE plots over time will be presented for ECG values by treatment group. Scatter plots of post-baseline versus baseline values will be presented by treatment group.

The frequency of subjects who have abnormal values will be summarized by treatment group and study period. All data for subjects that have abnormal values will be presented in a separate listing. The following abnormality ranges will be used for ECG parameters (Table 6):

Table 6: ECG Abnormality Ranges

	Low	Normal	High
HR (bpm)	<45	45-90	> 90
PR Interval (msec)	<120	120-220	>220
QRS Interval (msec)	-	<120	≥120
QT Interval (msec)	-	<500	≥500

Criteria for abnormal corrected QTc interval values and changes from baseline are presented in Table 7. Abnormal values and changes will be summarized using frequency distribution tables over time by treatment group and period. A listing of subjects meeting any of the abnormality criteria will be provided.

Table 7: Criteria for Abnormal QTc Values and Changes From Baseline

Parameter	Classification	Criteria
QTc value	Normal	≤450
	> 450 – 480	>450 - ≤480
	> 480 – 500	>480 – ≤500
	> 500	> 500
QTc change from baseline	No concern	≤30
	Concern	>30 – 60
	Clear concern	> 60

These criteria are based on ICH E14 Guideline

5.6. C-SSRS

The Columbia Suicide Severity Rating scale (C-SSRS) is a low-burden measure of the spectrum of suicidal ideation and behavior. It is a semi structured clinician-administered questionnaire designed to solicit the occurrence, severity, and frequency of suicide-related ideation and behaviors during the assessment period.

Using the C-SSRS, potentially suicide-related events will be categorized using the following scores:

Suicidal Ideation (1-5)

1: Wish to be dead

2: Non-specific active suicidal thoughts

3: Active suicidal ideation with any methods (not plan) without intent to act

4: Active suicidal ideation with some intent to act, without specific plan

5: Active suicidal ideation with specific plan and intent

Suicidal Behavior (6-10)

6: Preparatory acts or behavior

7: Aborted attempt

8: Interrupted attempt

9: Non-fatal suicide attempt

10: Completed suicide

A frequency distribution at each scheduled time point and endpoint will be provided by treatment group and study period. A listing will be provided of C-SSRS items throughout the study for subjects with suicidal ideation or behavior at any time point.

6. PHARMACOGENOMIC ANALYSES

DNA samples will be analyzed for the FAAH gene. The relationship between genetic subgroups for the FAAH gene (C/C or A/C) and JNJ-42165279 CCI) will be examined through descriptive statistics and graphical presentations. A similar analysis will be performed to explore the relationship between genetic subgroups and the CCI

7. BIOMARKERS

Biomarker analyses will be based on the biomarker analysis set.

During the study, the following biomarker evaluations will be performed at the start of the lead-in period, at Week 2 and at Week 6 of the double-blind treatment period and at Week 2 of the withdrawal period/end of withdrawal visit: plasma concentrations of FAAs (AEA, PEA and OEA) and blood and saliva collection for exploratory biomarkers. The analysis of FAAs is included in this SAP, a separate biomarker SAP will be written for the analysis of the exploratory biomarkers.

7.1. Summary of FAA Biomarker Values

The FAA biomarker analytes will be summarized CCI using descriptive statistics of actual values, changes and percent CCI Mean-SE plots CCI , and results will be listed.

Associations between CCI will be explored for each of the analytes using scatterplots and Pearson and Spearman correlation coefficients.

7.2. Relationship Between FAA Biomarker Analytes and CCI

To investigate the CCI biomarkers, associations between FAA levels at the CCI will be explored using scatterplots and Pearson and Spearman correlation coefficients.

Associations between FAA levels at the CCI CCI will be explored using scatterplots and Pearson and Spearman correlation coefficients.

7.3. Relationship Between FAA Biomarker Analytes and CCI

Descriptive statistics of actual values, changes and percent CCI will be provided CCI for all FAA biomarker analytes CCI for CCI

REFERENCES

1. Bech P, Gram LF, Dein E, Jacobsen O, Vitger J, Bolwig TG. Quantitative rating of depressive states. Acta Psychiatr Scand 1975; 51(3): 161-170.

2. Hamilton M. The assessment of anxiety states by rating. Br J Med Psychol 1959; 32(1): 50-55.

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

4. Lee S. et al: Does elimination of placebo responders in a placebo run-in increase the treatment effect in randomized clinical trials? A meta-analytic evaluation; Depression and Anxiety 2004;19(1)10-19

5. McIntyre A. Gendron A, McIntyre A. Quetiapine adjunct to selective serotonin reuptake inhibitors or venlafaxine in patients with major depression, comorbid anxiety, and residual depressive symptoms: A randomized, placebo-controlled pilot study. Depress Anxiety.2007;24(7):487-494