Janssen Research & Development

Statistical Analysis Plan

Double-Blind, Placebo-Controlled, Multi-Centre Study Investigating the Efficacy, Safety, and Tolerability of JNJ-61393215 as Adjunctive Treatment in Adults with Major Depressive Disorder with Anxious Distress with Suboptimal Response to Standard Antidepressants

Protocol 61393215MDD2001; Phase 2a

JNJ-61393215 (Orexin-1)
Amendment 2

Status:	Approved
Date:	30 September 2021
Prepared by:	Janssen Research & Development, LLC
Document No.:	EDMS-RIM-714672

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

TABLE OF CONTENTS

TABLE OF CONTENTS **2**

ABBREVIATIONS **4**

SAP AMENDMENT SUMMARY OF CHANGES TABLE **5**

1. INTRODUCTION **7**
1.1. Trial Objectives 7
1.2. Trial Design 9
1.3. Statistical Hypotheses for Trial Objectives 10
1.4. Sample Size Justification 10
1.5. Randomization and Blinding 10

2. GENERAL ANALYSIS DEFINITIONS **11**
2.1. Visit Windows 11
2.2. Pooling Algorithm for Analysis Centers 13
2.3. Analysis Sets 13
2.3.1. All Enrolled Analysis Set 13
2.3.2. All Randomized Analysis Set 13
2.3.3. Efficacy Analysis Sets 14
2.3.4. Safety Analysis Set 14
2.3.5. Pharmacokinetics Analysis Set 14
2.3.6. Immunogenicity Analysis Set 14
2.3.7. Pharmacodynamics Analysis Set 14
2.4. Definition of Subgroups 14
2.5. Study Day and Relative Day 14
2.6. Baseline 15
2.7. Imputation Rules for Missing AE Date/Time of Onset/Resolution 15

3. INTERIM ANALYSIS AND DATA MONITORING COMMITTEE REVIEW **16**

4. SUBJECT INFORMATION **16**
4.1. Demographics and Baseline Characteristics 16
4.2. Disposition Information 17
4.3. Treatment Compliance 17
4.4. Extent of Exposure 18
4.5. Protocol Deviations 18
4.6. Prior and Concomitant Medications 18

5. EFFICACY **18**
5.1. Analysis Specifications 18
5.1.1. Level of Significance 18
5.1.2. Data Handling Rules 19
5.2. Primary Efficacy Endpoint 19
5.2.1. Definition 19
5.2.2. Estimand 19
5.2.3. Analysis Methods 19
5.3. Major Secondary Endpoint 21
5.3.1. Definition 21
5.3.2. Analysis Methods 21
5.4. Other Efficacy Variables 21
5.4.1. Definition 21
5.4.2. Analysis Methods 22

6. SAFETY **23**

6.1. Adverse Events 23
6.2. Clinical Laboratory Tests 23
6.3. Vital Signs and Physical Examination Findings 24
6.4. Electrocardiogram 24
6.5. Other Safety Parameters 26
6.5.1. Columbia Suicide Severity Rating Scale (C-SSRS) 26

7. PHARMACOKINETICS/PHARMACODYNAMICS 27

8. BIOMARKERS 27

9. HEALTH ECONOMICS 27

ABBREVIATIONS

AE	adverse event
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
AR(1)	first order autoregressive process
AST	aspartate aminotransferase
BMI	body mass index
BP	blood pressure
CI	confidence interval
CRF	case report form
CSR	Clinical Study Report
C-SSRS	Columbia Suicide Severity Rating Scale
DB	double blind
DSM-5	Diagnostic and Statistical Manual of Mental Disorders-Fifth Edition
ECG	electrocardiogram
FAS	full analysis set
FU	follow-up
GAD-7	generalized anxiety disorder-7
HAM-A	Hamilton Anxiety Rating scale
$HDRS_{17}$	Hamilton Depression Rating Scale 17 items
IDS-C30	Inventory of Depressive Symptomatology, Clinician Rating -30
ITT	Intent-to-Treat
IWRS	interactive web response system
LS	least-squares
MDD	major depressive disorder
MedDRA	Medical Dictionary for Regulatory Activities
MINI	MINI International Neuropsychiatric Inventory
mITT	modified Intent-to-Treat
MMRM	mixed model for repeated measures
PD	pharmacodynamic(s)
PHQ-9	patient health questionnaire -9 item
PK	pharmacokinetic(s)
QTcB	QT interval corrected using Bazett's formula
QTcF	QT interval corrected using Fridericia's formula
SAE	serious adverse event
SAP	Statistical Analysis Plan
SD	standard deviation
SE	Sleep Efficiency
TEAE	treatment-emergent adverse event
US	United States
WHO-DD	World Health Organization Drug Dictionary

SAP AMENDMENT SUMMARY OF CHANGES TABLE

DOCUMENT HISTORY	
Document	**Date**
Amendment 2	30 September 2021
Amendment 1	18 January 2021
Original SAP	10 February 2020

Amendment 2

Section number and Name	**Description of Change**	**Brief Rationale**
Sections 2.3.3 (Efficacy Analysis Sets) and 5.2.3 (Analysis Methods)	Revising the primary analysis dataset by including patients who were ongoing at the time of the trial suspension due to the COVID-19 pandemic.	Received feedback from the FDA on the primary analysis set for protocol amendment 3: against excluding patients who were ongoing at the trial suspension due to Covid-19 from the primary analysis set– it violates the intent-to-treat principle and may limit interpretation of the study results.

Amendment 1

Section number and Name	**Description of Change**	**Brief Rationale**
1.1 Trials Objectives 2.1 Visit Windows 4.1 Demographics and Baseline Characteristics 5.4 Other Efficacy Variables	Remove objective, visit window, baseline characteristics and efficacy variable related to the PDSS-SR	PDSS-SR will not be performed in the study due to logistical reasons
1.4 Sample Size Justification	Increase the potential maximum sample size increasing from 15% to 25% to count for COVID-19 related dropouts	Take into account for potential dropouts, and exclusion of participants who were ongoing at the time of the trial suspension from the primary analyses due to COVID-19 which is unexpected, unforeseen and unprecedented circumstance.
2.3.3 Efficacy Analysis Sets 5.2.3 Analysis Method	Revise the primary analysis dataset by excluding participants who were ongoing at the time of the trial suspension due to the COVID-19 pandemic	Minimize potential confounding effect as the unexpected and unprecedented circumstances around the COVID-19 pandemic may be associated with increased anxiety and thus might impact primary outcome measures in the study population of participants with MDD with anxious distress.

Section number and Name	Description of Change	Brief Rationale
1.1 Trials Objectives 2.1 Visit Windows 4.1 Demographics and Baseline Characteristics 5.4 Other Efficacy Variables	Remove objective, visit window, baseline characteristics and efficacy variable related to the PDSS-SR	PDSS-SR will not be performed in the study due to logistical reasons
4.1. Demographics and Baseline Characteristics	Revised “full analysis dataset” to “all randomized analysis dataset”	To be consistent with other trials

1. INTRODUCTION

This statistical analysis plan (SAP) contains definitions of analysis sets, derived variables, and statistical methods for all planned analyses for the clinical study report (CSR) for Study 61393215MDD2001.

This SAP does not include planned analyses on biomarkers or pharmacogenomics data or on population pharmacokinetic (PK), pharmacodynamics (PD) and exposure/response analyses, which will be specified as appropriate in separate documents.

1.1. Trial Objectives

Objectives	Endpoints
Primary	
• To evaluate the efficacy of JNJ-61393215 as adjunctive treatment compared to adjunctive placebo, as assessed by the change from baseline to week 6 on a 17-item Hamilton Depression Rating Scale ($HDRS_{17}$) in participants with MDD with anxious distress with a score ≥ 2 on item 26 or 27 of the Inventory of Depressive Symptomatology, Clinician Rating -30 (IDS-C30), who have a suboptimal response to current treatment with a standard antidepressant.	• Change from baseline in $HDRS_{17}$ total score at Week 6
Key Secondary	
• To evaluate the impact of adjunctive treatment with JNJ-61393215 compared to adjunctive placebo on the severity of anxiety as measured by the change in the Hamilton Anxiety Rating scale (HAM-A) from baseline to week 6.	• Change from baseline in HAM-A total score at Week 6
Secondary	
Efficacy	
• To evaluate the impact of adjunctive treatment with JNJ-61393215 compared to adjunctive placebo on $HDRS_{17}$ from baseline to weeks 2 and 4	• Change from baseline in $HDRS_{17}$ total score at Weeks 2 and 4

Objectives	Endpoints
• To evaluate the impact of adjunctive treatment with JNJ-61393215 compared to adjunctive placebo on the severity of anxiety as measured by the change in HAM-A from baseline to weeks 2 and 4	• Change from baseline in HAM-A total score at Weeks 2 and 4
• To evaluate the efficacy of JNJ-61393215 as adjunctive treatment compared to adjunctive placebo, as assessed by the change from baseline to week 6 on the $HDRS_{17}$ in participants with a baseline HAM-A score ≥20	• Change from baseline in $HDRS_{17}$ total score at Week 6 in participants with a baseline HAM-A total score ≥20
• To evaluate the efficacy of JNJ-61393215 as adjunctive treatment compared to adjunctive placebo, as assessed by the change from baseline to Week 6 on the HAM-A in participants with a baseline HAM-A score ≥20	• Change from baseline in HAM-A total score at Week 6 in participants with a baseline HAM-A total score ≥20
• To evaluate the impact of adjunctive treatment with JNJ-61393215 compared to adjunctive placebo on participant-reported severity of anxiety as measured by the change in the General Anxiety Disorder-7 scale (GAD-7) from baseline to Week 6	• Change from baseline in GAD-7 total score at Week 6
• To evaluate the impact of adjunctive treatment with JNJ-61393215 compared to adjunctive placebo on participant-reported severity of symptoms of major depressive disorder (MDD), as measured by the change in the Patient Health Questionnaire (PHQ-9) from baseline to weeks 2, 4 and 6	• Change from baseline in PHQ-9 total score at Weeks 2, 4 and 6
Safety	
• To investigate the overall safety and tolerability of adjunctive treatment with JNJ-61393215 in participants with MDD with anxious distress	Safety assessments including: • Adverse events (AEs) • Proportion of all serious adverse events (SAEs) • Vital signs, physical examinations, electrocardiogram (ECG), and laboratory parameters • Columbia Suicide Severity Rating Scale (C-SSRS)

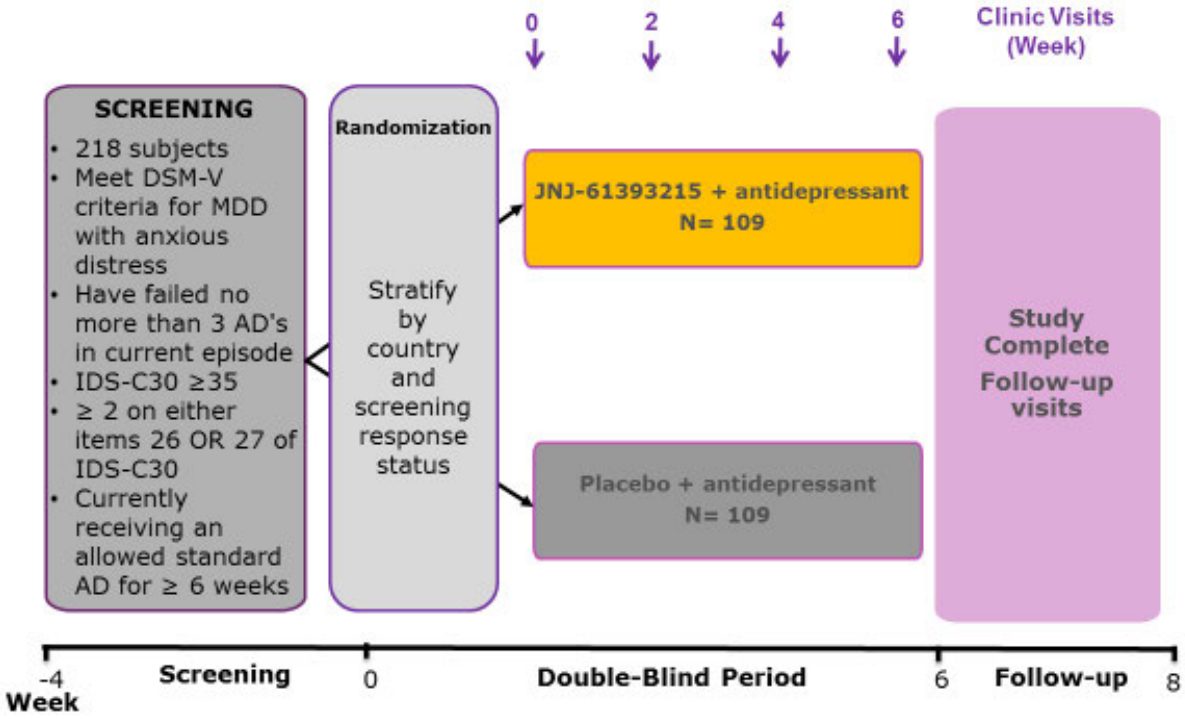
1.2. Trial Design

This is a double-blind (DB), placebo-controlled, randomized, parallel-group, multicenter study in participants with MDD with anxious distress. An estimated target of approximately 218 participants will be randomly assigned in this study with 109 participants planned per treatment group. The randomization will be stratified by country and screening response status (whether participants have improvement in percent changes in $HDRS_{17}$ total score $\geq 25\%$ at baseline from screening or $HDRS_{17}$ total score < 18 at baseline). The target study population includes male and female participants, between 18 and 64 years of age inclusive, with a Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5) diagnosis of MDD with anxious distress, confirmed by the Mini International Neuropsychiatric Inventory (MINI) Plus with module for MDD with anxious distress

The study will consist of 3 phases: a screening phase of up to 4 weeks, a double-blind treatment phase of 6 weeks, and a posttreatment follow-up (FU) phase of 2 weeks. The total duration of participation will be approximately 12 weeks for each participant.

A schematic overview of the study design is provided in Figure 1.

Figure 1: Schematic Overview of the Study Design



1.3. Statistical Hypotheses for Trial Objectives

The primary hypothesis of this study is that treatment with JNJ-61393215 will lead to significant improvement in depressive symptoms from baseline compared to placebo when administered as an adjunctive treatment to a standard antidepressant in the treatment of MDD patients with anxious distress with a score ≥ 2 on item 26 or 27 of the IDS-C30 who have a suboptimal response to current standard treatment. Significant improvement will be demonstrated by improvement in depressive symptoms from baseline to the 6-week endpoint in the $HDRS_{17}$ total score.

1.4. Sample Size Justification

The estimated sample size of 218 participants (109 participants per group) was determined based on the assumption of an effect size of at least 0.4 for the $HDRS_{17}$ total score (mean change from baseline to Week 6 endpoint between the JNJ-61393215 and placebo groups of 3 units with standard deviation (SD) = 7.5). This is considered to be a clinically relevant difference in a population with suboptimal response to standard oral antidepressant therapy. The SD of 7.5 in the change in $HDRS_{17}$ total score from baseline is a reasonable assumption based on previously conducted clinical trials in a similar patient population (40411813DAX2001 and 42165279MDD2001). Power is set at 90%, with a 1-sided alpha level of 0.10 and a 6-week drop--out rate of 10%.

It was also assumed that 15% of the randomized participants would be excluded from the primary efficacy analysis due to having improvement in percent changes in $HDRS_{17}$ total score ≥ 25% at baseline from screening or $HDRS_{17}$ total score < 18 at baseline. The estimated participants to be included in the primary analysis is 166. A blinded data review for purpose of sample size re-estimation may be performed. Sample size may be re-adjusted if the observed screening response rate, the drop-out rate (including dropouts due to the COVID-19 pandemic) or the standard deviation (SD) substantially deviate from the assumed value. Sample size can increase by a maximal 25%, so that the potential maximal number of participants to be enrolled in this trial would be 272.

1.5. Randomization and Blinding

Central randomization will be implemented in this study. Participants will be randomly assigned to 1 of 2 intervention groups based on a computer-generated randomization schedule prepared before the study by or under the supervision of the sponsor. The randomization will be stratified by country and screening response status (whether participants have improvement in percent changes in $HDRS_{17}$ total score ≥ 25% at baseline from screening or $HDRS_{17}$ total score < 18 at baseline). The criteria for screening response will be blinded for the investigator. The interactive web response system (IWRS) will assign a unique intervention code, which will dictate the intervention assignment and matching study intervention kit for the participant. The requestor must use his or her own user identification and personal identification number when contacting the IWRS and will then give the relevant participant details to uniquely identify the participant.

In general, randomization codes will be disclosed fully only if the study is completed and the clinical database is closed.

2. GENERAL ANALYSIS DEFINITIONS

2.1. 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 below are the visit windows and the target days for each visit. The reference day is Study Day 1 (which is the first day the study drug was taken in the double-blind phase). 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.

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. Listed below are the visit windows and the target days for each visit defined in the protocol.

Assessments prior to Day 1 of the double-blind phase that are not considered a baseline assessment will be labeled as 'Screening'.

For subjects who have multiple follow-up visits, the scheduled visit will be used in the summaries, however, data for all follow-up visits will be displayed in listings when applicable.

If there is an unscheduled assessment with collection date that doesn't fall into any visit window, then it will be assigned to a visit window that is closest to its collection date, and if this unscheduled assessment has a collection date equidistant from the dates for two adjacent visit windows, then it will be assigned to the visit window that is latter.

Table 1: Visit Window

Parameter	Analysis Phase	Scheduled Visit	Time Interval (label on output)	Time Interval (Day)	Target Time Point (Day/Time)
$HDRS_{17}$, HAM-A	Screening	Screening	Screening	< 1	-28 to -24
	DB	Baseline	Baseline	< 1	Day 1
		Week 2	Week 2	Visit Week 2 (as collected on eCRF)	Day 15
		Week 4	Week 4	Visit Week 4 (as collected on eCRF)	Day 29
		Week 6	Week 6	Visit Week 6 (as collected on eCRF)	Day 43
	FU	Follow-up	Follow-up	FU Visit (as collected on eCRF)	Day 57
PHQ-9	DB	Baseline	Baseline	< 1	Day 1
		Week 2	Week 2	Visit Week 2 (as collected on eCRF)	Day 15
		Week 4	Week 4	Visit Week 4 (as collected on eCRF)	Day 29
		Week 6	Week 6	Visit Week 6 (as collected on eCRF)	Day 43
	FU	Follow-up	Follow-up	FU Visit (as collected on eCRF)	Day 57
GAD-7	DB	Baseline	Baseline	< 1	Day 1
		Week 6	Week 6	Visit Week 6 (as collected on eCRF)	Day 43
ECG	Screening	Screening	Screening	<1	-28 to -24
	DB	Day 1[a]	Day 1	< 1	Day 1
		Week 4	Week 4	Visit Week 4 (as collected on eCRF)	Day 29
		Week 6	Week 6	Visit Week 6 (as collected on eCRF)	Day 43
	FU	Follow-up	Follow-up	FU Visit (as collected on eCRF)	Day 57
Physical Exam, Vital Signs, Oral or Tympanic Temperature	Screening	Screening	Screening	<1	-28 to -24
	DB	Baseline	Baseline	< 1	Day 1
		Week 4	Week 4	Visit Week 4 (as collected on eCRF)	Day 29
	FU	Follow-up	Follow-up	FU Visit (as collected on eCRF)	Day 57
C-SSRS	Screening	Screening	Screening	<1	-28 to -24

Parameter	Analysis Phase	Scheduled Visit	Time Interval (label on output)	Time Interval (Day)	Target Time Point (Day/Time)
	DB	Baseline	Baseline	< 1	Day 1
		Week 2	Week 2	Visit Week 2 (as collected on eCRF)	Day 15
		Week 4	Week 4	Visit Week 4 (as collected on eCRF)	Day 29
		Week 6	Week 6	Visit Week 6 (as collected on eCRF)	Day 43
	FU	Follow-up	Follow-up	FU Visit (as collected on eCRF)	Day 57
Labs (Hematology, Chemistry, Urinalysis)	Screening	Screening	Screening	<1	-28 to -24
	DB	Baseline	Baseline	< 1	Day 1
		Week 2	Week 2	Visit Week 2 (as collected on eCRF)	Day 15
		Week 4	Week 4	Visit Week 4 (as collected on eCRF)	Day 29
		Week 6	Week 6	Visit Week 6 (as collected on eCRF)	Day 43
	FU	Follow-up	Follow-up	FU Visit (as collected on eCRF)	Day 57

a: 2 hours postdose

2.2. Pooling Algorithm for Analysis Centers

Subjects will be enrolled at sites in the USA, United Kingdom, Russia, Ukraine and Moldova. Actual site enrollment rates will be monitored to avoid gross imbalances across centers. Country or sites will not be pooled for analyses. To account for country variability, country will be used as a factor in the statistical models to analyze efficacy.

2.3. Analysis Sets

Subjects will be classified into the following analysis sets.

2.3.1. All Enrolled Analysis Set

The all enrolled analysis set includes all subjects who were not screen failures and were assigned to a treatment group.

2.3.2. All Randomized Analysis Set

The all randomized analysis set includes all subjects who were randomized in the study.

2.3.3. Efficacy Analysis Sets

The full analysis set (FAS) includes all randomized subjects who received at least 1 dose of study agent and have both a baseline and at least 1 postbaseline efficacy assessment.

The enriched analysis set (EAS) includes subjects in FAS with having improvement in percent changes in $HDRS_{17}$ total score < 25% at baseline from screening and $HDRS_{17}$ total score ≥ 18 at baseline.

2.3.4. Safety Analysis Set

The safety analysis set includes all randomized subjects who received at least 1 dose of study agent.

2.3.5. Pharmacokinetics Analysis Set

Not applicable (NA). Analyses for pharmacokinetic data will be conducted by the PK group and will be provided in a separate document.

2.3.6. Immunogenicity Analysis Set

NA as there is no immunogenicity data for this trial.

2.3.7. Pharmacodynamics Analysis Set

NA as there is no pharmacodynamics data for this trial.

2.4. Definition of Subgroups

Subgroup	Definition
Screening response status	• Improvement in percent changes in $HDRS_{17}$ total score < 25% at baseline from screening and $HDRS_{17}$ total score ≥ 18 at baseline. • Improvement in percent changes in HDRS17 total score ≥ 25% at baseline from screening or HDRS17 total score < 18 at baseline
Baseline HAM-A status	• Baseline HAM-A total score ≥ 20 • Baseline HAM-A total score < 20

Other subgroups may also be explored.

2.5. Study Day and Relative Day

Study day or relative day for a visit is defined as:

- Visit date - (date of Study Day 1) +1, if visit date is ≥ date of Day 1
- Visit date - date of Day 1, if visit date < date of Day 1

There is no 'Day 0'.

2.6. Baseline

Baseline is defined as the last observation prior to the start of the first study agent administration.

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

Partial AE onset dates will be imputed as follows:

- If the onset date of an AE is missing day only, it will be set to:
  - First day of the month that the AE occurred, if month/year of the onset of AE is different than the month/year of the study agent start
  - The day of study agent start, if the month/year of the onset of AE is the same as month/year of the study agent start date and month/year of the AE resolution date is different
  - The day of study agent start or day of AE resolution date, whichever is earliest, if month/year of the onset of AE and month/year of the study agent start date and month/year of the AE resolution date are same
- If the onset date of an AE is missing both day and month, it will be set to the earliest of:
  - January 1 of the year of onset, as long as this date is on or after the study agent start date
  - Month and day of the study agent start date, if this date is the same year that the AE occurred
  - Last day of the year if the year of the AE onset is prior to the year of the study agent start date,
  - The AE resolution date.
- Completely missing onset dates will not be imputed.

Partial AE resolution dates not marked as ongoing will be imputed as follows:

- If the resolution date of an AE is missing day only, it will be set to the earliest of the last day of the month of occurrence of resolution or the day of the date of death, if the death occurred in that month.
- If the resolution date of an AE is missing both day and month, it will be set to the earliest of December 31 of the year or the day and month of the date of death, if the death occurred in that year.
- Completely missing resolution dates will not be imputed.

AE onset/resolution dates with missing times will be imputed as follows:

- A missing time of onset of an adverse event will be set to the earlier of:
  - 00:01 as long as the onset date is after the study agent start date
  - The time of the study agent start if this is the same day the AE occurred.
- The missing time of resolution of an adverse event will be set to 23:59.

If a missing time is associated with a partial or missing date, the date will be imputed first prior to imputing the time.

3. INTERIM ANALYSIS AND DATA MONITORING COMMITTEE REVIEW

NA as no IA and DMC are planned. However, a blinded data review for purpose of sample size re-estimation may be performed. Sample size may be re-adjusted if the observed screening response rate, the drop-out rate or the standard deviation (SD) substantially deviate from the assumed value. Sample size can increase with maximal 25%.

4. SUBJECT INFORMATION

The number of subjects in each analysis set will be summarized and listed by treatment group, and overall. In addition, the distribution of subjects by country, and site ID will be presented.

4.1. Demographics and Baseline Characteristics

Table 2 presents a list of the demographic variables that will be summarized by treatment group and overall for the all randomized analysis set.

Table 3 presents a list of the baseline psychiatric history variables that will be summarized by treatment group and overall for the all randomized analysis set.

Table 2: Demographic Variables

Continuous Variables:	**Summary Type**
Age (years), calculated based on date of informed consent date	Descriptive statistics (N, mean, standard deviation [SD], median and range [minimum and maximum])
Baseline weight (kg)	
Baseline height (cm)	
Baseline BMI (kg/m^2) calculated as Weight (kg)/[Height (m)]^2	
Categorical Variables:	
Sex (male, female, undifferentiated)	Frequency distribution with the number and percentage of subjects in each category
Race[a] (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or other Pacific Islander, White, Multiple)	
Ethnicity (Hispanic or Latino, not Hispanic or Latino)	

a If multiple race categories are indicated, the Race is recorded as 'Multiple'

Table 3: Psychiatric History Variables

Continuous Variables:	**Summary Type**
Baseline $HDRS_{17}$ total score	Descriptive statistics (N, mean, standard deviation [SD], median and range [minimum and maximum])
Baseline HAM-A total score	
Baseline GAD-7 total score	
PHQ-9 total score	
IDS-C30 total score	
Categorical Variables:	

Baseline HAM-A total score (≥ 20, <20)	Frequency distribution with the number and percentage of subjects in each category
Screening response status ("1" = defined as improvement in percent changes in $HDRS_{17}$ total score ≥ 25% at baseline from screening or $HDRS_{17}$ total score < 18 at baseline; else "2")	
Anxious distress status ("Yes" if a score ≥ 2 on item 26 or 27 of the Inventory of Depressive Symptomatology; else "No")	

4.2. Disposition Information

Screened subjects and reason for screen failures will be summarized overall.

The number of subjects in the following disposition categories will be summarized throughout the study by treatment group and overall:

- Subjects randomized
- Subjects receiving study agent
- Subjects completing the study
- Subjects who discontinued study agent
- Reasons for discontinuation of study agent
- Subjects who terminated study prematurely
- Reasons for termination of study

Listings of subjects will be provided for the following categories:

- Subjects who discontinued study agent
- Subjects who terminated study prematurely
- Subjects who were unblinded during the study period
- Subjects who were randomized yet did not receive study agent.

4.3. Treatment Compliance

The primary method to capture the treatment compliance will be based on AiCure or the combination of eCRF and AiCure if the exposure data from AiCure is insufficient.

In addition, study agent compliance that from eCRF will be calculated as follows as secondary method:

Compliance (%) = (number of capsules taken / number of capsules supposed to have been taken)*100.

Study agent compliance will be summarized descriptively.

Descriptive statistics on the percent (%) compliance will be summarized by treatment group for the safety analysis set. In addition, percent compliance will be categorized as <60%, 60%-<80%, 80%-100%, ≥100% and the number and percentage of subjects in each category will be summarized.

4.4. Extent of Exposure

The number and percentage of subjects who receive study agent will be summarized by treatment group.

Descriptive statistics for study agent duration (N, mean, SD, median, and range (minimum, maximum)) will be presented by treatment group for the safety analysis set.

4.5. Protocol Deviations

Subjects with major protocol deviations will be identified prior to database lock and the subjects with major protocol deviations will be summarized by category.

4.6. Prior and Concomitant Medications

Prior and Concomitant medications will be coded using the World Health Organization Drug Dictionary (WHO-DD). Prior medications are defined as any therapy used before the day of first dose (partial or complete) of study agent. Concomitant medications are defined as any therapy used on or after the same day as the first dose of study agent, including those that started before and continue on after the first dose of study agent.

Summaries of concomitant medications will be presented by ATC term, treatment group, and study phase (DB phase and FU phase). The proportion of subjects who receive each concomitant medication will be summarized as well as the proportion of subjects who receive at least 1 concomitant medication.

Prior medications will be summarized by treatment group and ATC term.

5. EFFICACY

Primary analyses will focus on the EAS. However, all analyses will be done on enriched and full analysis set.

5.1. Analysis Specifications

5.1.1. Level of Significance

Unless otherwise specified, all statistical comparisons will be carried out at a 1-sided significance level of 10% and confidence intervals will be presented at 2-sided confidence level of 80% (i.e. equivalent to 1-sided confidence level of 90%).

5.1.2. Data Handling Rules

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

5.2. Primary Efficacy Endpoint

5.2.1. Definition

The HDRS17 is a clinician-administered rating scale designed to assess the severity of symptoms in subjects diagnosed with depression with a score range of 0 to 52. Each of the 17 items is rated by the clinician on either a 3- or a 5-point scale. The HDRS17 consists of 17 items that cover all of the core depressive symptoms (depressed mood, feeling of guilt, suicide, insomnia early, insomnia middle, insomnia late, work and activities, retardation [psychomotor], agitation, anxiety [psychological], anxiety somatic, somatic symptoms general, genital symptoms, hypochondriasis, loss of weight and insight). A total score (0 to 52) is calculated by adding the scores of all 17 items. For each item as well as the total score, a higher score represents a more severe condition. If 2 or more items are missing, no imputation will be performed, and the total score will be left missing. Otherwise, the total score will be calculated as sum of the nonmissing items multiplied by the ratio of the maximum number of items (ie, 17) to the number of nonmissing items.

5.2.2. Estimand

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

Population: participants with MDD with anxious distress with suboptimal response to standard antidepressants.

Variable: change from baseline to Week 6 in the $HDRS_{17}$ total score.

Intervention event: the effect of the initially randomized treatment that would have been observed had all participants remained on their treatment throughout the double-blind treatment phase.

Population-level summary: the difference in mean change from baseline to Week 6 post dose in the $HDRS_{17}$ total score total score between treatment conditions.

5.2.3. Analysis Methods

The primary efficacy analysis will be based on the EAS. The primary comparison will be between Orexin-1 and placebo using the primary efficacy endpoint, i.e., change in $HDRS_{17}$ total score from baseline to Week 6.

The JNJ-61393215 treatment group will be compared with the placebo group using the primary efficacy endpoint, changes from baseline in $HDRS_{17}$ total score, with the comparison performed by means of a mixed-effects model using repeated measures (MMRM), with time, treatment

(placebo, JNJ-61393215), country, and time-by-treatment interaction as factors, baseline $HDRS_{17}$ total score as a continuous covariate. An unstructured variance-covariance matrix will be used. In case of convergence problems, alternative variance-covariance structures will be tried in the following order, with the first structure that converges being used in the analysis: heterogeneous Toeplitz, standard Toeplitz, and AR(1) (first order autoregressive process) with separate subject random effect. The comparison of JNJ-61393215 versus placebo will be performed using the appropriate contrasts. The contrast on Week 6 changes will be of primary interest and tested at one-sided alpha level of 0.10. Time profile of means for values and changes from baseline in $HDRS_{17}$ total score will be represented graphically over time by treatment group.

The analyses for changes from baseline in $HDRS_{17}$ total score will also be carried out in the full analysis set using a mixed-effects model using repeated measures (MMRM), with time, treatment (placebo, JNJ-61393215), country, screening response status (using $HDRS_{17}$) and time-by-treatment interaction as factors, and baseline $HDRS_{17}$ total score as a continuous covariate. An unstructured variance-covariance matrix will be used. In case of convergence problems, alternative variance-covariance structures will be tried in the following order, with the first structure that converges being used in the analysis: heterogeneous Toeplitz, standard Toeplitz, and AR(1) with separate subject random effect.

In addition, sensitivity analyses with the same MMRM analysis by excluding participants who were ongoing at the time of the trial suspension due to the COVID-19 pandemic and/or by excluding noncompliance subjects will be conducted.

Sensitivity Analysis for Missing Data

The sensitivity analysis will be based on an analysis of covariance (ANCOVA) model using change from baseline to Week 6 based on LOCF data. The models will include factors for treatment, country and screening response status, and baseline $HDRS_{17}$ total score as a continuous covariate.

If the overall missingness of the primary endpoint at Week 6 is above 15%, then additional sensitivity analysis, the delta adjustment multiple imputation method, that rely on Missing Not at Random assumptions may be performed to assess the robustness of the primary analysis and MMRM results. This method consists the following 3 steps:

Step 1 – Multiple imputation
If there are subjects with non-monotone missing data pattern, datasets with only monotone missing data patterns will be created first by imputing the intermediate missing values using the Markov Chin Monte Carlo method. If all subjects have a monotone missing data pattern, the MAR-based multiple imputation with the regression will be used to impute missing values.

Step 2 – Delta adjustments
The imputed values for subjects who discontinued will be adjusted by adding δ_P to the imputed values for subjects in placebo group and adding δ_A to the imputed values for subjects in JNJ61393215 group.

Step 3 – Analysis
Same MMRM as described for the primary efficacy analysis will be conducted.

Subgroup Analysis

Analysis on changes in $HDRS_{17}$ total score from baseline to Week 6 in baseline HAM-A status subgroup (whether subjects had a baseline HAM-A total score ≥20) will also be conducted using descriptive statistics and the similar MMRM method that is aforementioned for the primary end point.

5.3. Major Secondary Endpoint

5.3.1. Definition

The HAM-A measures the severity of a patient's anxiety, based on 14 parameters, including anxious mood, tension, fears, insomnia, somatic complaints and behavior at the interview. The major value of HAM-A is to document the results of pharmaco- or psychotherapy, rather than as a diagnostic or screening tool. Each item is simply given a 5-point score - 0 (not present) to 4 (severe). In this study, the Structured Interview Guide version of the HAM-A (SIGH-A) will be used. A total score (0 to 56) is calculated by adding the scores of all 14 items. For each item as well as the total score, a higher score represents a more severe condition.

5.3.2. Analysis Methods

The analyses for changes from baseline in HAM-A total score at weeks 2, 4 and 6 will be carried out using a mixed-effects model using repeated measures (MMRM), with time, treatment (placebo, JNJ-61393215), country, screening response status (using HDRS17) and time-by-treatment interaction as factors, and baseline HAM-A total score as a continuous covariate. An unstructured variance-covariance matrix will be used. In case of convergence problems, alternative variance-covariance structures will be tried in the following order, with the first structure that converges being used in the analysis: heterogeneous Toeplitz, standard Toeplitz, and AR(1) with separate subject random effect. The comparison of JNJ-61393215 versus placebo will be performed using the appropriate contrasts. The contrast on Week 6 changes will be of primary interest and tested at one-sided alpha level of 0.10. Time profile of means for values and changes from baseline in HAM-A total score will be represented graphically over time by treatment group.

In addition, analysis on changes in HAM-A total score from baseline to Week 6 in baseline HAM-A status subgroup (whether subjects had a baseline HAM-A total score ≥20) will also be conducted using the similar MMRM method that is aforementioned.

5.4. Other Efficacy Variables

5.4.1. Definition

GAD-7
The GAD-7 is a self-reported questionnaire for screening and severity measuring of generalized anxiety disorder (GAD). GAD-7 has seven items, which measure severity of various signs of GAD according to reported response categories with assigned points (see below). Assessment is

indicated by the total score, which made up by adding together the scores for the scale all seven items. The scale uses a normative system of scoring as shown below, with question at the end qualitatively describing severity of the patient's anxiety over the past 2 weeks.

- Not at all (0 points)
- Several days (1 point)
- More than half the days (2 points)
- Nearly every day (3 points)

A total score (0 to 21) is calculated by adding the scores of all 7 items. For each item as well as the total score, a higher score represents a more severe condition.

PHQ-9

The PHQ-9 is a participant-reported measure of depressive symptomatology. The PHQ-9 is a 9-item scale, where each item is rated on a 4-point scale (0=Not at all, 1=Several Days, 2=More than half the days, and 3=Nearly every day), with a total score range of 0 to 27. The recall period is 2 weeks. For each item as well as the total score, a higher score represents a more severe condition.

5.4.2. Analysis Methods

All analyses will be done on enriched and full analysis set.

GAD-7

The analyses for changes from baseline in GAD-7 total score at week 6 will be carried out in the FAS using an analysis of covariance (ANCOVA) model, with factors for treatment (placebo, JNJ-61393215), country, screening response status (using $HDRS_{17}$) and baseline GAD-7 total score as a continuous covariate. The treatment effects will be estimated using least squares means.

PHQ-9

The analyses for changes from baseline in PHQ-9 total score at weeks 2, 4 and 6 will be carried out in the FAS using a mixed-effects model using repeated measures (MMRM), with time, treatment (placebo, JNJ-61393215), country, screening response status (using HDRS17) and time-by-treatment interaction as factors, and baseline PHQ-9 total score as a continuous covariate. An unstructured variance-covariance matrix will be used. In case of convergence problems, alternative variance-covariance structures will be tried in the following order, with the first structure that converges being used in the analysis: heterogeneous Toeplitz, standard Toeplitz, and AR(1) with separate subject random effect. The comparison of JNJ-61393215 versus placebo will be performed using the appropriate contrasts. The contrast on Week 6 changes will be of primary interest and tested at one-sided alpha level of 0.10. Time profile of means for values and changes from baseline in PHQ-9 total score will be represented graphically over time by treatment group.

6. SAFETY

All safety analyses will be based on the safety analysis set based on actual treatment received, unless otherwise specified.

For all continuous safety variables, descriptive statistics will include the N, mean, standard deviation, median, minimum, and maximum. Categorical variables will be summarized using frequency counts and percentages.

6.1. Adverse Events

The verbatim terms used in the CRF by investigators to identify adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). Any AE occurring at or after the initial administration of study agent in double blind phase is considered to be treatment emergent. If the event occurs on the day of the initial administration of study agent, and either event time or time of administration are missing, then the event will be assumed to be treatment emergent. If the event date is recorded as partial or completely missing, then the event will be considered to be treatment emergent unless it is known to be prior to the first administration of study agent based on partial onset date or resolution date. All reported treatment-emergent adverse events will be included in the analysis. For each adverse event, the number and percentage of subjects who experience at least 1 occurrence of the given event will be summarized by treatment group.

Summary tables will be provided for treatment-emergent adverse events:

- AEs
- Serious AEs (SAEs)
- AEs leading to discontinuation of study agent
- AEs by severity
- AEs by relationship to study agent

In addition to the summary tables, listings will be provided for subjects who:

- Had SAEs
- Had AEs leading to discontinuation of study agent

6.2. Clinical Laboratory Tests

Clinical laboratory tests will be displayed for the subjects included in the safety analysis set.

Descriptive statistics will be presented for chemistry, hematology, and urinalysis laboratory tests at scheduled time points and displayed by treatment group.

Graphical displays will be presented for aspartate aminotransferase (AST) and alanine aminotransferase (ALT) laboratory tests at scheduled time points.

Change from baseline to scheduled time points will be summarized for chemistry, hematology, and urinalysis tests and displayed by treatment group.

Shift tables will be provided summarizing the shift in laboratory values from baseline to scheduled time points with respect to abnormality criteria (low, normal, high).

6.3. Vital Signs and Physical Examination Findings

Continuous vital sign parameters including temperature, weight, pulse, blood pressure (systolic and diastolic), and Body Mass Index (BMI) will be summarized at each assessment time point. Body Mass Index will be calculated as weight (kg)/(height (m))^2. The weight and height measurements collected at screening will be used in the calculation. Change from baseline will be summarized for the treatment. Descriptive statistics (mean, standard deviation, median, minimum and maximum) will be presented.

Incidence of treatment-emergent clinically important abnormalities in vital signs while on treatment, as defined in Table 5, will be summarized for subjects who had a baseline assessment and at least 1 postbaseline assessment for that vital sign. If the baseline value is missing, the postbaseline value will be compared against the abnormally low/abnormally high criteria.

A listing of subjects with treatment-emergent clinically important abnormalities in vital signs will be presented, along with a listing of all vital sign measurements.

Table 5: Clinically Important/Markedly Abnormal Vital Signs

Vital Sign	Criteria
Pulse	>[120] bpm and with >[30] bpm increase from baseline
	<[50] bpm and with >[20] bpm decrease from baseline
Systolic blood pressure	>[180] mm Hg and with >[40] mm Hg increase from baseline
	<[90] mm Hg and with >[30] mm Hg decrease from baseline
Diastolic blood pressure	>[105] mm Hg and with >[30] mm Hg increase from baseline
	<[50] mm Hg and with >[20] mm Hg decrease from baseline
Temperature	>[38]oC and with ≥[1]oC increase from baseline

In addition, a by-subject listing of the abnormal physical examination data will be presented.

6.4. Electrocardiogram

The ECG parameters that will be analyzed are heart rate, PR interval, QRS interval, QT interval, and corrected QT (QTc) interval using the following correction methods: Bazett's formula (QTcB), and Fridericia's formula (QTcF).

Bazett's formula:
$QTcB\ (msec) = QT\ (msec) * (HR(bpm)/60)^{0.5}$;

Fridericia's formula:
$QTcF\ (msec) = QT\ (msec) * (HR(bpm)/60)^{0.33}$

The number and percentage of subjects with QTc interval will be summarized at each scheduled time point. The number and percentage of subjects with QTc interval increases from baseline to the maximum postbaseline value will be summarized. Refer to the following table for summary categories.

Criteria for Abnormal QTc Values and Changes From Baseline	
QTc value	≤450
	>450 – 480
	>480 – 500
	>500
QTc change from baseline	≤30
	>30 – ≤60
	> 60

A shift table will be provided summarizing the shift from baseline to maximum QTc interval classification at each scheduled time point.

Descriptive statistics of ECG parameters and change from baseline will be summarized at each scheduled time point.

If ECG measurements are repeated at a visit, they will be averaged. The averaged value will be considered the 'Visit' ECG result.

A postbaseline abnormality (abnormality based on criteria defined below) will be considered treatment emergent (TE) if it is worse than the baseline abnormality. If the baseline abnormality is missing, the abnormality is always considered treatment emergent. If the postbaseline value is above the upper limit and the baseline value is below the lower limit, then the postbaseline abnormality will be considered TE. The same applies to the postbaseline value being below the lower limit with the baseline value being above the upper limit.

The number and percentage of subjects with treatment-emergent ECG values outside predefined limits/abnormal postbaseline values (relative to baseline) will be presented by treatment group over time for each phase of the study:

- Heart rate (bpm): [<50] and [>100]
- PR interval (msec): [<120] and [>200]
- QRS interval (msec): [>120]
- QTc (msec): [>500]

The interpretation of the ECGs as determined by a qualified physician (investigator or qualified designee) will be displayed by the number and percentage of subjects meeting the normality criteria. The interpretation will be summarized over time by treatment phase.

Listings will be produced for all ECG data including unscheduled visit data. A listing of clinically relevant ECG abnormalities will also be provided.

6.5. Other Safety Parameters

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

The 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: Actual Attempt (non-fatal)
10: Completed Suicide

If no events qualify for a score of 1 to 10, a score of 0 will be assigned (0="no event that can be assessed on the basis of C-SSRS"). Higher scores indicate greater severity.

A frequency distribution at each time point by treatment group will be provided. Shifts from baseline to the maximum postbaseline score during the double-blind phase will be summarized by treatment group.

The maximum postbaseline score during the double-blind phase assigned for each subject will be summarized into one of three broad categories: No suicidal ideation or behavior (0), Suicidal ideation (1-5), Suicidal behavior (6-10). Shifts from baseline to the maximum category during the double-blind phase will be summarized by treatment group.

In addition, a frequency distribution at each time point for whether the "subject has engaged in non-suicidal self-injurious behavior" will be presented by treatment group.

A listing of C-SSRS items throughout the study for subjects with Suicidal Ideation or Behavior at any time point will be provided.

7. PHARMACOKINETICS/PHARMACODYNAMICS

The pharmacokinetics analyses will be conducted by the PK group and the details will be provided in a separate document.

Analyses on pharmacodynamics are not applicable as no pharmacodynamics data for the trial.

8. BIOMARKERS

Details of the biomarker analyses will be provided in a separate SAP.

9. HEALTH ECONOMICS

Analyses on health economics are not applicable as no health economics data for the trial.

Janssen Research & Development

Statistical Analysis Plan
Amendment 2

Double-Blind, Placebo-Controlled, Multi-Centre Study Investigating the Efficacy, Safety, and Tolerability of JNJ-61393215 as Adjunctive Treatment in Adults with Major Depressive Disorder with Anxious Distress with Suboptimal Response to Standard Antidepressants

Protocol 61393215MDD2001; Phase 2a

JNJ-61393215 (Orexin-1)

Author: PPD

Approved by:

PPD

Digitally signed by PPD
DN: cn=PPD, o=Janssen Pharma, ou=Experimental Medicine, email=PPD, c=BE
Reason: I am approving this document.
Date: 2021.09.30 17:01:53 +02'00'
Adobe Acrobat version: 11.0.20

PPD, MD Date
Clinical Leader

PPD

Digitally signed by PPD
DN: c=US, o=JNJ, ou=Subscribers, cn=PPD PP, 0.9.2342.19200300.100.1.1=10011182
Reason: I am approving this document.
Date: 2021.09.30 18:15:53 +02'00'
Adobe Reader version: 11.0.20

PPD MD Date
ED Statistical Functional Manager

NOTE: The final SAP must be signed-off before database lock.