Janssen Research & Development *

Statistical Analysis Plan

An Open-label, Long-term, Safety and Efficacy Study of Intranasal Esketamine in Treatment-resistant Depression

Safety and Sustenance of Esketamine Treatment Response With Repeated Doses at Intervals Determined by Symptom Severity (SUSTAIN-2)

Protocol ESKETINTRD3004; Phase 3

JNJ-54135419 (esketamine)

Status:	Approved
Date:	10 November 2017
Prepared by:	Janssen Research & Development, LLC
Document No.:	EDMS-ERI-131851138, 1.0

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

TABLE OF CONTENTS

LIST OF IN-TEXT TABLES

LIST OF IN-TEXT FIGURES

ABBREVIATIONS

AD	antidepressant
AE	adverse event
BMI	body mass index
BP	Blood Pressure
BPIC-SS	Bladder Pain / Interstitial Cystitis Symptom Score
BPRS+	Four-item positive symptom subscale of the brief psychiatric rating scale
CADSS	Clinician Administered Dissociative States Scale
CGADR	Clinical Global Assessment of Discharge Readiness
CGI-S	Clinical Global Impression – Severity
CRF	case report form
C-SSRS	Columbia-Suicide Severity Rating Scale
DB	Double-blind
D/C	Discontinued
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders (4th edition)
ECG	Electrocardiogram
eCRF	electronic case report form
EQ-5D-5L	EuroQol Group; 5 dimension; 5 level
EQ-VAS	EuroQol Group: visual analogue scale
EU	European Union
FDA	Food and Drug Administration
FU	Follow-Up
GAD-7	Generalized Anxiety Disorder 7-item scale
HVLT-R	Hopkins Verbal Learning Test-Revised
IDMC	Independent Data Monitoring Committee
ICH	International Conference on Harmonization
IDS-C_{30}	Inventory of Depressive Symptoms-Clinician rated, 30 item
LOCF	last observation carried forward
MADRS	Montgomery-Asberg Depression Rating Scale
MDD	major depressive disorder
MedDRA	Medical Dictionary for Regulatory Activities
MINI	Mini International Neuropsychiatric Interview
MGH-ATRQ	Massachusetts General Hospital Antidepressant Treatment History
MOAA/S	Modified observer's assessment of alertness/sedation
OL	Open-label
PHQ-9	Patient health questionnaire – 9
PWC-20	Physician Withdrawal Checklist; 20 item
SAE	serious adverse event
SAP	Statistical Analysis Plan
SD	standard deviation
SDS	Sheehan Disability Scale
SE	Standard error
SOC	System Organ Class
SNRI	Serotonin and Norepinephrine Reuptake Inhibitors
SSRI	Selective Serotonin Reuptake Inhibitors
TEAEs	Treatment-emergent adverse events
TEMA	treatment-emergent markedly abnormal
TRD	Treatment Resistant Depression
XR	Extended release

1. INTRODUCTION

This statistical analysis plan (SAP) contains definitions of analysis sets, derived variables, and statistical methods for all planned analyses for study JNJ54135419-ESKETINTRD3004.

1.1. Trial Objectives

Primary Objective

The primary objective of this study is to assess the long-term safety and tolerability of intranasal esketamine plus a newly initiated oral antidepressant in subjects with TRD, with special attention to the following:

- Potential effects on cognitive function
- Potential treatment-emergent symptoms of cystitis and/or lower urinary tract symptoms
- Potential withdrawal and/or rebound symptoms following cessation of intranasal esketamine treatment

Secondary Objectives

To assess the effect of intranasal esketamine plus a newly initiated oral antidepressant in subjects with TRD on

- Safety and tolerability with special attention to the following:

  Treatment-emergent adverse events (TEAEs), including TEAEs of special interest

  Local nasal tolerability

  Effects on heart rate, blood pressure, respiratory rate and blood oxygen saturation

  Effects on alertness and sedation

  Potential psychosis-like effects

  Dissociative symptoms

  Potential effects on suicidal ideation/behavior.

- Long-term efficacy, including effects on:

  Depressive symptoms (clinician and self-reported), overall severity of depressive illness, functional impairment and associated disability, anxiety symptoms, and health-related quality of life and health status

- Response rate over time, defined as:
  - percentage of subjects with ≥50% reduction from baseline (induction phase) in the Montgomery-Asberg Depression Rating Scale (MADRS) total score,
  - percentage of subjects with ≥50% reduction from baseline (induction phase) in the Patient Health Questionnaire, 9-item (PHQ-9) total score

- Remission rate over time, defined as:
  - percentage of subjects with MADRS total score ≤12,
  - percentage of subjects with PHQ-9 total score ≤5

Exploratory Objectives

- To assess the potential relationship of biomarkers with response/non-response to intranasal esketamine plus an oral antidepressant in subjects with TRD
- To assess medical resource utilization
- To assess subject tradeoff preferences for key benefit and harm outcomes associated with TRD treatment, using a stated-choice conjoint analysis survey. Reporting of these survey results may be conducted separately from this study.

1.2. Trial Design

This is an open-label, multicenter, long-term study to evaluate the safety and efficacy of intranasal esketamine plus a newly initiated oral antidepressant in subjects with TRD. Subjects will enter the study either directly (referred to as 'direct-entry subjects') or after completing the double-blind induction phase of ESKETINTRD3005, a short-term efficacy study in elderly subjects with TRD (referred to as 'transferred-entry subjects'). Approximately 750 direct entry subjects will be enrolled in this study, plus transferred-entry subjects from study ESKETINTRD3005. A total of at least 100 subjects 65 years or older (who are either direct entry subjects or transferred entry subjects from the ESKETINTRD3005 study) will be enrolled.

ESKETINTRD3005 is a randomized, double-blind, active-controlled, 4-week study in male and female elderly subjects (≥65 years) with TRD to assess the efficacy, safety, and tolerability of flexibly dosed intranasal esketamine (28 mg, 56 mg or 84 mg) plus a newly initiated oral antidepressant, compared with a newly initiated oral antidepressant (active comparator) plus intranasal placebo.

Transferred-entry subjects who are non-responders (defined as <50% reduction in the MADRS total score from baseline [Day 1] at the end of the 4-week double-blind induction phase of ESKETINTRD3005 study) will be referred to as 'transferred-entry non-responder subjects' in the subsequent sections of the SAP. Transferred-entry subjects who are responders (defined as ≥50% reduction in the MADRS total score from baseline [Day 1]) at the end of the 4-week double-blind induction phase of ESKETINTRD3005 study) will be referred to as 'transferred-entry responder subjects' in the subsequent sections of the SAP.

An Independent Data Monitoring Committee (IDMC) will be commissioned for this study to review safety data periodically.

This study (ESKETINTRD3004) has 4 phases:

- Up to 4-week screening phase (direct-entry subjects only)

- A 4-week open-label induction phase (direct-entry subjects and transferred-entry non-responder subjects)
- A 48-week open-label optimization/maintenance phase (all responder subjects from the open-label induction phase of the current study, and transferred-entry responder subjects).
- A 4-week follow-up phase (for all subjects treated with intranasal esketamine)

The maximum duration of the subject's participation in ESKETINTRD3004 study will be 60 weeks for direct-entry subjects; 56 weeks for transferred-entry non-responder subjects, and 52 weeks for transferred-entry responder subjects. The end of the study will occur when at least 300 subjects have received treatment with intranasal esketamine for 6 months and at least 100 subjects for 12 months (Note: the total number of subjects will be based on subjects from this study and subjects from other intranasal esketamine Phase 3 studies).

A description of the study phases is provided below.

Screening Phase

After giving informed consent, direct-entry subjects with TRD who are ≥18 years of age (or older if the minimum legal age of consent in the country in which the study is taking place is >18 years), will be screened to determine eligibility for study participation. Transferred-entry subjects from ESKETINTRD3005 study will not participate in this 4-week screening phase.

Direct-entry subjects will be eligible for screening regardless of whether or not they are currently taking oral antidepressant treatments. At screening, subjects must have had a nonresponse to ≥2 different oral antidepressant treatments in the current episode of depression, as assessed by the Massachusetts General Hospital-Antidepressant Treatment Response Questionnaire (MGH-ATRQ), and confirmed by documented records (eg, medical/pharmacy/prescription records or a letter from treating physician etc,), in their current episode of depression.

Eligible subjects taking antidepressant medication(s) at the start of the screening phase must discontinue all of their current oral antidepressant medication(s), being used for depression treatment, including adjunctive/augmentation therapies, prior to the start of the induction phase. Of note, subjects taking benzodiazepines (at dosages equal to or less than the equivalent of 6 mg/day of lorazepam) and/or permitted non-benzodiazepine sleep medications (eg, zolpidem, zaleplon) during the screening phase can continue these medications. No dose increases beyond the equivalent of 6 mg/day of lorazepam, or new benzodiazepine medications are permitted. If clinically indicated, the antidepressant treatments may either be tapered and discontinued during the screening phase, or, discontinued and switched directly to 1 of the 4 new oral antidepressant medication(s) on Day 1 of the open-label induction phase, per clinical judgment.

Subjects not currently taking oral antidepressant medication(s) at screening will start 1 of the 4 new oral antidepressant medication (duloxetine, escitalopram, sertraline, or venlafaxine extended release [XR]) on Day 1 of the open-label induction phase.

Subjects meeting the inclusion/exclusion criteria are eligible to proceed to the open-label induction phase.

Open-label Induction Phase (4 weeks)

Direct-entry subjects and transferred-entry non-responder subjects will participate in this phase.

Direct-entry subjects: Intranasal esketamine treatment will be self-administered at scheduled treatment sessions in the clinic/study site, twice weekly for 4 weeks. Subjects who are <65 years old will start intranasal esketamine with an initial dose of 56 mg on Day 1, with the dose adjusted based on efficacy and tolerability in the subsequent visits of the induction phase (flexible dose: 56 mg or 84 mg). Subjects who are ≥65 years old will start intranasal esketamine with an initial dose of 28 mg on Day 1, with the dose adjusted based on efficacy and tolerability (28, 56 or 84 mg) in the subsequent visits of the induction phase. In addition, all direct-entry subjects will initiate a new, open-label oral antidepressant on Day 1, which should be taken daily during the study.

The assigned oral antidepressant will be 1 of 4 oral antidepressant medications (duloxetine, escitalopram, sertraline, or venlafaxine extended release [XR]), that the subject has not previously had a nonresponse to in the current depressive episode, has not been previously intolerant to (lifetime), and is available in the participating country.

Transferred-entry non-responder subjects from ESKETINTRD3005 study (all will be ≥65 years old) will join this study at the start of the open-label induction phase. These subjects will start intranasal esketamine with an initial dose of 28 mg on Day 1, with the dose adjusted based on efficacy and tolerability (28, 56 or 84 mg) in the subsequent visits of the induction phase. These subjects should continue taking the same oral antidepressant (at the same dose) during the study, as taken in the last week of the double-blind induction phase of ESKETINTRD3005 study.

Transferred-entry subjects' may participate in this study only if it is clinically appropriate in the opinion of the investigator.

For transferred-entry non-responder subjects, results of all assessments performed on Day 28 of the induction phase of that study (Visit 2.9 of ESKETINTRD3005 study) will not be repeated as part of Visit 2.1 of the current study. For these subjects, the Day 28 visit of the ESKETINTRD3005 study will coincide with Day 1 (Visit 2.1) for the current study. There is no gap allowed between studies.

Optimization/Maintenance Phase:

Responder subjects at the end of the induction phase (≥50% reduction in MADRS from baseline [Day 1] to end of induction phase) of the ESKETINTRD3004 study, will be eligible to proceed to the optimization/maintenance phase; and continue receiving open-label intranasal esketamine treatment (at the same dose; 28 mg, 56 mg, or 84 mg) and should continue to take the same oral antidepressant (at the same dose) as taken in the last week of the induction phase of ESKETINTRD3004 study, unless poorly tolerated, in which case the oral antidepressant may be discontinued after review with sponsor.

Non-responders at the end of the induction phase of the current study will complete an early withdrawal visit and proceed to the follow-up phase.

Eligible transferred-entry responder subjects from the ESKETINTRD3005 study (all will be ≥65 years old) will join the current study starting from the optimization/maintenance phase. Transferred-entry subjects may participate in this study only if, it is clinically appropriate in the opinion of the investigator. These subjects will start intranasal esketamine with an initial dose of 28 mg (Week 5; Study Day 32) and have their dose adjusted over the following 3 weeks of the optimization/maintenance phase. Subjects should continue to take the same oral antidepressant (at the same dose) during the study, as taken in the last week of the double-blind induction phase of ESKETINTRD3005 study, unless poorly tolerated, in which case the oral antidepressant may be discontinued after review with sponsor.

For transferred-entry responder subjects, results of all assessments performed on Day 28 of the induction phase (Visit 2.9 of ESKETINTRD3005 study) will not be repeated as part of Visit 3.1 of the current study. The Day 28 visit of the ESKETINTRD3005 study will coincide with Day 28 (Visit 3.1) for the current study. There is no gap allowed between studies.

For all subjects, the intranasal treatment session frequency will be reduced from that in the induction phase (twice weekly) to weekly for the first 4 weeks of optimization/maintenance phase (ie, Week 5 to Week 8). After the first 4 weeks, the frequency of intranasal treatment sessions will be adjusted to either once weekly or once every other week based on the severity of depressive symptoms, as assessed by the MADRS total score. A maximum of 3 changes in intranasal treatment session frequency from weekly to every other week is permitted during the optimization/maintenance phase.

Follow-up Phase:

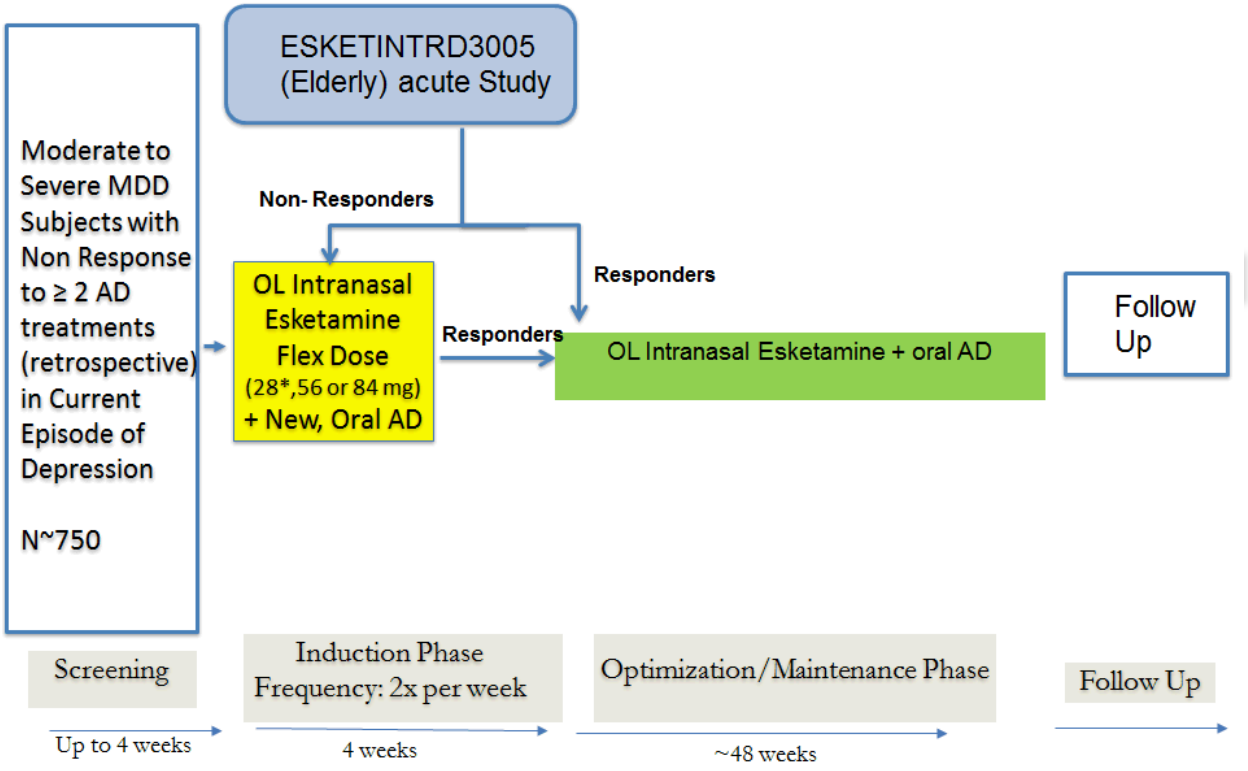
This phase will include all subjects who have received at least 1 dose of intranasal study medication in this study. Follow-up visits will be performed at 1, 2 and 4 weeks after the last dose of intranasal study drug.

At the start of the follow-up phase, further clinical/standard of care for the treatment of depression will be arranged by the study investigator and/or the subject's treating physician. There will be no intranasal esketamine administered during this phase. Subject will be provided with an additional 4 week supply of the oral antidepressant medication, to ensure that there is no interruption of oral antidepressant therapy during the transition to further clinical/standard of care.

The decision to continue the oral antidepressant in this phase will be at the discretion of the investigator. However, in order to better assess potential withdrawal symptoms from intranasal study drug and facilitate maintenance of clinical benefit, the oral antidepressant medication should be continued for the duration of the follow-up phase unless determined as not clinically appropriate.

A diagram of the study design is provided in Figure 1.

Figure 1: Schematic Overview of the Study



* Intranasal esketamine 28 mg dose is the only option for subjects ≥65 years.
Note: At entry to the ESKETINTRD3004, transferred entry non-responder subjects will continue to receive the same oral antidepressant initiated in the ESKETINTRD3005 study. The new oral AD is for direct entry subjects only
AD, antidepressant; MDD, major depressive disorder; OL, open-label;

An additional open-label safety extension study, 54135419TRD3008, may be available (pending country and site approval) for eligible subjects participating in the ESKETINTRD3004 study. Refer to the 54135419TRD3008 protocol for full details.

1.3. Statistical Hypotheses for Trial Objectives

There is no formal hypothesis for this safety study.

1.4. Sample Size Justification

No formal sample size calculation was performed for the study. The projected sample size of 750 direct entry subjects plus transferred entry subjects is considered adequate to obtain at least 300 subjects who have received treatment with esketamine for 6 months and at least 100 subjects for 12 months (Note: the total number of subjects will be based on subjects from this study and subjects from other intranasal esketamine Phase 3 studies). The number of transferred entry subjects is based on predictions related to discontinuation rate and efficacy, and therefore may vary.

The sample size will include at least 100 older subjects (who are either direct entry subjects or transferred entry subjects from the ESKETINTRD3005 study) aged ≥65 years.

1.5. Randomization and Blinding

This is an open label study. Randomization and blinding procedures will not be applicable for this study. All the subjects will self-administer intranasal esketamine and an oral antidepressant in an open label manner.

However, the blind of intranasal treatment from ESKETINTRD3005 study will not be broken for transferred-entry subjects. Since there may be responders to intranasal placebo and oral antidepressant who did not receive intranasal esketamine prior to ESKETINTRD3004 study, all transferred-entry responder subjects will start at a dose of 28 mg of intranasal esketamine.

2. GENERAL ANALYSIS DEFINITIONS

2.1. Analysis Phases

This study will have 3 analysis phases:

- A 4-week open-label induction phase (direct-entry subjects and transferred-entry non-responders)
- A 48-week open-label optimization/maintenance phase (all responder subjects from the open label induction phase of the current study and transferred-entry responder subjects)
- A 4-week follow-up phase (post-treatment phase)
- Each analysis phase has its own analysis phase start and end dates.

2.1.1. Study Reference Start and End Dates

For direct-entry and transferred-entry non-responder subjects, the reference start date for the study is defined as the earlier date of the first dose of intranasal study drug or the oral antidepressant (the date is missing for screened subjects who did take not any intranasal study drug or oral antidepressant). For transferred-entry responder subjects, the reference start date is the date of first dose of oral antidepressant medication in the optimization/maintenance phase. The reference end date for subjects who have completed/discontinued from the study is the maximum of the date of last visit and trial completion/discontinuation.

2.1.2. Analysis Phase Start and End Dates

Screening Phase

The start date of the screening phase is the informed consent date. For subjects who continue to the open-label induction phase, the screening phase end date is the open-label induction phase start date.

The start and end dates for the screening phase are defined for direct entry subjects only.

Open-Label Induction Phase (direct-entry and transferred-entry non-responder subjects)

The start date of the open-label induction phase (referred to as, 'IND start date') is the earlier of date of the first dose of intranasal study medication or oral antidepressant in the open-label induction phase. For subjects who enter the optimization/maintenance phase, the open-label induction phase end date (referred to as, 'IND end date') is the date of first dose of oral antidepressant taken in the optimization/maintenance phase or the date of the last visit in the open-label induction phase if date of the first oral medication is missing. For subjects who discontinue in the open-label induction phase, IND end date is the maximum of date of the last visit in the open-label induction phase or date of early termination from the induction phase.

The start date/time of the open-label induction phase (referred to as, 'IND start date/time') is the IND start date and the time of the first dose of intranasal study medication in this phase. If no intranasal study medication is administered or it is administered after the start of oral antidepressant, then the time will be left blank.

The start and end dates for the open-label induction phase are defined for direct entry subjects and transferred-entry non-responders only.

Optimization/Maintenance Phase

The start date of the optimization/maintenance phase (referred to as, 'OP/MA start date') is the date of the first dose of oral antidepressant medication in the optimization/maintenance phase. For subjects who complete/discontinue from the optimization/maintenance phase, the phase end date (referred to as, 'OP/MA end date') is the maximum of date of last visit in the optimization/maintenance phase or the date of completion/withdrawal from the optimization/maintenance phase.

Follow-Up Phase

The start date of the follow-up (post-treatment) phase (referred to as, 'F/U start date') is the day after the end date of the last treatment phase in which the subject participated. For subjects who complete/discontinue from the follow-up phase, the follow-up phase end date (referred to as, 'F/U end date') is the maximum of the last follow-up visit date or end of trial date.

2.1.3. Study Day and Relative Day

Study day is calculated relative to the reference start date for the study. Relative day is calculated relative to the analysis phase start date of the analysis phase in which the data are captured. A minus (-) sign indicates days prior to the start of study or prior to the start of the analysis phase.

Study day for an event on or after the start of the study is calculated as:

event date- reference start date+1.

Study day for an event prior to the start of the study is calculated as:

event date-reference start date

Relative day for an event on or after the analysis phase start date is calculated as:

event date – analysis phase start date + 1.

Relative day for an event prior to the analysis phase start date is calculated as:

event date – analysis phase start date.

There is no study day 0 or relative day 0.

2.2. Baseline and End Point

Baseline is defined for each parameter/assessment.

- Open-label Induction phase:
  - For direct-entry and transferred-entry non-responder subjects: The last observation prior to or on the start date of open-label induction phase is denoted as, 'Baseline (IND)'. For transfer-entry responder subjects: Baseline (IND) is from study 3005. Baseline (IND) will be copied from the 3005 study into the 3004 study.
  - Baseline for the ECG measurements is defined as the average of all non-missing assessments on or before the date of the first dose of study medication for direct-entry and transferred-entry responders. Baseline for transferred-entry non-responders is the last observation prior to or on the start date of the open-label induction phase. The 'Average Predose' is computed in the database for study 3005 and will be copied into study 3004.
- Optimization/Maintenance phase: The last observation prior to or on the start date of the optimization/maintenance phase is denoted as 'Baseline (OP/MA)'. For most subjects this pertains to assessments performed on Day 28 of the open-label induction phase for direct-entry and transferred-entry non-responder subjects (Visit 2.9), and Day 28 of the double-blind induction phase for transferred-entry responder subjects (Visit 2.10 of study ESKETINTRD3005) and is denoted as 'Baseline (OP/MA)'.

For each variable measured over time, the 'End Point (IND)' value is defined as the last postbaseline assessment value during the open-label induction phase for direct-entry subjects and transfer-entry non-responder subjects. This value will be the same as the Baseline (OP/MA) value for subjects who continue into the optimization/maintenance phase from the open-label induction phase of 3004.

The 'End Point (OP/MA)' value is defined as the last postbaseline assessment value during the optimization/maintenance phase.

2.3. Visit Windows

As subjects do not always adhere to the protocol visit schedule, the following rules are applied to assign actual visits to protocol 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 that any study drug was taken in the open-label induction phase for direct-entry subject and transferred-entry non-responders, and the first day any study drug was taken in the optimization/maintenance phase for responders from induction phase and transferred-entry responders).

If a subject has 2 or more scheduled or unscheduled visits in one visit window, the visit closest to the target day will be used as the protocol visit for that visit window. 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 is assigned to a visit window, it will no longer be used for a later time point except for the end point.

Listed below are the visit windows and the target days (if applicable) for each visit defined in the protocol for all phases (Table 1).

Table 1: Analysis Visits

Parameter	Analysis Phase	Scheduled Day	Time Interval (label on output)	Time Interval[a] (Day)	Target Time Point from start of each phase (Day)
MADRS	Screening	Within 4 Weeks of Day 1	Screening		
	IND	1	Baseline (IND)	<= 1	1
		8	Day 8 (IND)	2-11	8
		15	Day 15 (IND)	12-18	15
		22	Day 22 (IND)	19-24	22
		28	Day 28 (IND)	25 to end of IND	28
		IND final visit	End Point (IND)	2 to end of IND	
	OP/MA Weekly x=1, 2, 3, ... , 46	28	Baseline (OP/MA)	1	1
		32	Week 1 (OP/MA)	2-8	5
		32 + x*7	Week (x+1) (OP/MA)	2 + x*7 to 8 + x*7	5 + x*7
		361	Week 48 (OP/MA)	331 to end of OP/MA	334
		OP/MA final visit	End Point (OP/MA)	2 to end of OP/MA	
	F/U	F/U Day 7 (Last dose +7 days)	F/U Week 1	1 to 10	7
		F/U Day 14 (Last dose +14 days)	F/U Week 2	11 to 21	14
		F/U Day 28 (Last dose +28 days)	F/U Week 4	22 to end of F/U	28
		F/U final visit	End Point (F/U)	1 to end of F/U	
PHQ-9 and EQ-5D-5L EQ-VAS and Health Status Index	Screening		Screening		
	IND	1	Baseline (IND)	<= 1	1
		15	Day 15 (IND)	2-21	15
		28	Day 28 (IND)	22 to end of IND	28
		IND final visit	End Point (IND)	2 to end of IND	
	OP/MA Every Two Weeks x=1, 2, 3, … , 21	28	Baseline (OP/MA)	1	1
		32	Week 1 (OP/MA)	2-12	5
		46	Week 3 (OP/MA)	13-26	19
		46 + x*14	Week (3 + x*2) (OP/MA)	13 + x*14 to 26 + x*14	19 + x*14
		354	Week 47 (OP/MA)	321 to end of OP/MA	327
		OP/MA final visit	End Point (OP/MA)	2 to end of OP/MA	
	F/U	F/U Day 14 (Last dose +14 days)	F/U Week 2	1 to 21	14
		F/U Day 28 (Last dose +28 days)	F/U Week 4	22 to end of F/U	28
		F/U final visit	End Point (F/U)	1 to end of F/U	

Table 1: Analysis Visits

Parameter	Analysis Phase	Scheduled Day	Time Interval (label on output)	Time Interval[a] (Day)	Target Time Point from start of each phase (Day)
CGI-S	Screening		Screening		
	IND	1	Baseline (IND)	<=1	1
		4	Day 4 (IND)	2-6	4
		8	Day 8 (IND)	7-9	8
		11	Day 11 (IND)	10-13	11
		15	Day 15 (IND)	14-18	15
		22	Day 22 (IND)	19-24	22
		28	Day 28 (IND)	25 to end of IND	28
		IND final visit	End Point (IND)	2 to end of IND	
	OP/MA Weekly x=1, 2, 3, ... , 46	28	Baseline (OP/MA)	1	1
		32	Week 1 (OP/MA)	2-8	5
		32 + x*7	Week (x+1) (OP/MA)	2 + x*7 to 8 + x*7	5 + x*7
		361	Week 48 (OP/MA)	331 to end of OP/MA	334
		OP/MA final visit	End Point (OP/MA)	2 to end of OP/MA	
	F/U	F/U Day 14 (Last dose +14 days)	F/U Week 2	1 to 21	14
		F/U Day 28 (Last dose +28 days)	F/U Week 4	22 to end of F/U	28
		F/U final visit	End Point (F/U)	1 to end of F/U	
GAD-7	Screening		Screening		
	IND	1	Baseline (IND)	<= 1	1
		28	Day 28 (IND)	2 to end of IND	28
		IND final visit	End Point (IND)	2 to end of IND	
	OP/MA Every 4 weeks x=1, 2, 3, … , 10	28	Baseline (OP/MA)	1	1
		53	Week 4 (OP/MA)	2-40	26
		53 + x*28	Week (4 + x*4) (OP/MA)	13 + x*28 to 40 + x*28	26 + x*28
		361	Week 48 (OP/MA)	321 to end of OP/MA	334
		OP/MA final visit	End Point (OP/MA)	2 to end of OP/MA	
	F/U	F/U Day 28 (Last dose +28 days)	F/U Week 4	1 to end of F/U	28
		F/U final visit	End Point (F/U)	1 to end of F/U	
SDS	Screening		Screening		
	IND	1	Baseline (IND)	<= 1	1
		15	Day 15 (IND)	2-21	15
		28	Day 28 (IND)	22 to end of IND	28
		IND final visit	End Point (IND)	2 to end of IND	
	OP/MA Every 4 weeks x=1, 2, 3, … , 10	28	Baseline (OP/MA)	1	1
		53	Week 4 (OP/MA)	2-40	26
		53 + x*28	Week (4 + x*4) (OP/MA)	13 + x*28 to 40 + x*28	26 + x*28
		361	Week 48 (OP/MA)	321 to end of OP/MA	334
		OP/MA final visit	End Point (OP/MA)	2 to end of OP/MA	

Table 1: Analysis Visits

Parameter	Analysis Phase	Scheduled Day	Time Interval (label on output)	Time Interval[a] (Day)	Target Time Point from start of each phase (Day)
	F/U	F/U Day 28 (Last dose +28 days)	F/U Week 4	1 to end of F/U	28
		F/U final visit	End Point (F/U)	1 to end of F/U	
Hematology, Chemistry, Urinalysis from LAB	Screening[b]		Screening		
	IND	1	Baseline (IND)	<= 1	1
		28	Day 28 (IND)	2 to end of IND	28
		IND final visit	End Point (IND)	2 to end of IND	
	OP/MA Every 8 weeks x=1, 2, 3	28	Baseline (OP/MA)	<=1	1
		53	Week 4 (OP/MA)	2-40	26
		109	Week 12 (OP/MA)	41-110	82
		109+56*x	Week (12+8*x) (OP/MA)	55+56*x to 110+56*x	82+56*x
		333	Week 44(OP/MA)	279-320	306
		361	Week 48(OP/MA)	321 to end of OP/MA	334
		OP/MA final visit	End Point (OP/MA)	2 to end of OP/MA	
	F/U	F/U Day 28 (Last dose +28 days)	F/U Week 4	1 to end of F/U	28
Urinalysis	IND	1	Baseline (IND)	<=1	1
		15	Day 15 (IND)	2-21	15
		28	Day 28 (IND)	22 to end of IND	28
		IND final visit	End Point (IND)	2 to end of IND	
	OP/MA Every 8 weeks x=1,2,3	28	Baseline (OP/MA)	<=1	1
		53	Week 4 (OP/MA)	2-40	26
		109	Week 12 (OP/MA)	41-110	82
		109+56*x	Week (12+8*x) (OP/MA)	55+56*x to 110+56*x	82+56*x
		333	Week 44 (OP/MA)	279-320	306
		361	Week 48(OP/MA)	321 to end of OP/MA	334
		OP/MA final visit	End Point (OP/MA)	2 to end of OP/MA	
	F/U	F/U Day 14 (Last dose +14 days)	F/U Week 2	1-21	14
		F/U Day 14 (Last dose +28 days)	F/U Week 4	22 to end of F/U	28
		F/U final visit	End Point(F/U)	1 to end of F/U	
Vital Signs: (TEMP [predose at each visit], BP[c], Pulse, RR (at each visit, predose, 40M, 1H, 1.5H)	IND	1	Baseline(IND)g Day 1 (IND): 40M Day 1 (IND): 1H Day 1 (IND): 1H30M	<=1/predose	1
		4	Day 4 (IND): Predose Day 4 (IND): 40M Day 4 (IND): 1H Day 4 (IND): 1H30M	2-6	4
		8	Day 8 (IND): Predose	7-9	8

Table 1: Analysis Visits

Parameter	Analysis Phase	Scheduled Day	Time Interval (label on output)	Time Interval[a] (Day)	Target Time Point from start of each phase (Day)
			Day 8 (IND): 40M Day 8 (IND): 1H Day 8 (IND): 1H30M		
		11	Day 11 (IND): Predose Day 11 (IND): 40M Day 11 (IND): 1H Day 11 (IND): 1H30M	10-13	11
		15	Day 15 (IND): Predose Day 15 (IND): 40M Day 15 (IND): 1H Day 15 (IND): 1H30M	14-16	15
		18	Day 18 (IND): Predose Day 18: 40M Day 18: 1H Day 18: 1H30M	17-20	18
		22	Day 22 (IND): Predose Day 22 (IND): 40M Day 22 (IND): 1H Day 22 (IND): 1H30M	21- 23	22
		25	Day 25 (IND): Predose Day 25 (IND): 40M Day 25 (IND): 1H Day 25 (IND): 1H30M	24 to end of IND	25
		IND final visit	End Point (IND)	Day 1: 40M to end of IND	
	OP/MA Weekly x=1, 2, 3, ... , 46	25	Baseline (OP/MA)	<=1	1
		32	Week 1 (OP/MA): Predose Week 1 (OP/MA): 40M Week 1 (OP/MA): 1H Week 1 (OP/MA): 1H30M	1-8	5
		32 + x*7	Week (x+1) (OP/MA): Predose Week (x+1) (OP/MA): 40M Week (x+1) (OP/MA): 1H Week (x+1)	2 + x*7 to 8 + x*7	5 + x*7

Table 1: Analysis Visits

Parameter	Analysis Phase	Scheduled Day	Time Interval (label on output)	Time Interval[a] (Day)	Target Time Point from start of each phase (Day)
			(OP/MA): 1H30M		
		361	Week 48 (OP/MA): Predose Week 48 (OP/MA): 40M Week 48 (OP/MA): 1H Week 48 (OP/MA): 1H30M	331 to end of OP/MA	334
		OP/MA final visit	End Point (OP/MA)	Day 1: 40M to end of OP/MA	
	F/U	F/U Day 14 (Last dose +14 days)	F/U Week 2	1 to 21	14
		F/U Day 28 (Last dose +28 days)	F/U Week 4	22 to end of F/U	28
		F/U final visit	End Point (F/U)	1 to end of F/U	
Weight, BMI, Nasal Examination	Screening		Screening		
	IND	1	Day 1	<= 1/predose	1
		28	Day 28 (IND)	2 to end of IND	28
		IND final visit	End Point (IND)	2 to end of IND	
	OP/MA Every 4 weeks x=1, 2, 3, … , 10	28	Baseline (OP/MA)	<=1	1
		53	Week 4 (OP/MA)	2-40	26
		53 + x*28	Week (4 + x*4) (OP/MA)	13 + x*28 to 40 + x*28	26 + x*28
		361	Week 48 (OP/MA)	321 to end of OP/MA	334
		OP/MA final visit	End Point (OP/MA)	2 to end of OP/MA	
	F/U (not for Weight)	F/U Day 28 (Last dose +28 days)	F/U Week 4	1 to end of F/U	28
		F/U final visit	End Point (F/U)	1 to end of F/U	
12-lead ECG	IND	Screening, 1	Baseline (IND)	<=1/predose	1
		1	Day 1 (IND): 1H	1	1
		8	Day 8 (IND): 1H	2-11	8
		15	Day 15 (IND): 1H	12-20	15
		25	Day 25 (IND): 1H	21 to end of IND	25
		IND final visit	End Point (IND)	Day 1: 1H to end of IND	
	OP/MA Every 4 weeks x=1, 2, 3, … , 10	25	Baseline (OP/MA)	<=1	1
		32	Week 1 (OP/MA): Predose	2-8	5
		32	Week 1 (OP/MA): 1H	2-8	5
		39	Week 2 (OP/MA): 1H	9-19	12

Table 1: Analysis Visits

Parameter	Analysis Phase	Scheduled Day	Time Interval (label on output)	Time Interval[a] (Day)	Target Time Point from start of each phase (Day)
		53	Week 4 (OP/MA): 1H	20-40	26
		53 + x*28	Week (4 + x*4) (OP/MA): 1H	13 + x*28 to 40 + x*28	26 + x*28
		361	Week 48 (OP/MA): 1H	321 to end of OP/MA	334
		OP/MA final visit	End Point (OP/MA)	2 to end of OP/MA	
	F/U	F/U Day 28 (Last dose +28 days)	F/U Week 4)	1 to end of F/U	28
Nasal Symptom Questionnaire (Predose and 1H postdose)	IND	1	Day 1 (IND)	<=1	1
		4	Day 4 (IND)	2-7	4
		11	Day 11 (IND)	8-14	11
		18	Day 18 (IND)	15-21	18
		25	Day 25 (IND)	22 to end of IND	25
		IND final visit	End Point (IND)	2 to end of IND	
	OP/MA Every 4 weeks x=1, 2, 3, … , 10	32	Week 1 (OP/MA)	1-15	5
		53	Week 4 (OP/MA)	16-40	26
		53 + x*28	Week (4 + x*4) (OP/MA)	13 + x*28 to 40 + x*28	26 + x*28
		361	Week 48 (OP/MA)	321 to end of OP/MA	334
		OP/MA final visit	End Point (OP/MA)	2 to end of OP/MA	
C-SSRS (BL/SC version and since last visit version)	Screening		Screening	<=1	
	IND	1	Baseline (IND)	<=1	1
		4	Day 4 (IND)	2-6	4
		8	Day 8 (IND)	7-9	8
		11	Day 11 (IND)	10-13	11
		15	Day 15 (IND)	14-16	15
		18	Day 18 (IND)	17-20	18
		22	Day 22 (IND)	21- 23	22
		25	Day 25 (IND)	24-26	25
		28	Day 28 (IND)	27 to end of IND	28
		IND final visit	End Point (IND)	2 to end of IND	
	OP/MA Weekly x=1, 2, 3, ... , 46	28	Baseline (OP/MA)	<=1	1
		32	Week 1 (OP/MA)	2-8	5
		32 + x*7	Week (x+1) (OP/MA)	2 + x*7 to 8 + x*7	5 + x*7
		361	Week 48 (OP/MA)	331 to end of OP/MA	334
		OP/MA final visit	End Point (OP/MA)	2 to end of OP/MA	
	F/U	F/U Day 14 (Last dose +14 days)	F/U Week 2	1 to 21	14
		F/U Day 28 (Last dose +28 days)	F/U Week 4	22 to end of F/U	28
		F/U final visit	End Point (F/U)	1 to end of F/U	
MOAA/S[d],	IND	1	Baseline (IND): only	<=1	1

Table 1: Analysis Visits

Parameter	Analysis Phase	Scheduled Day	Time Interval (label on output)	Time Interval[a] (Day)	Target Time Point from start of each phase (Day)
Pulse Oximetry[e] (predose and every 15 minutes to 1.5H), BPRS+ and CADSS (predose, 40M, 1.5H) CGADR[f], (1H, 1.5H)			for Pulse Oximetry		
		1	Day 1 (IND)	<=1	1
		4	Day 4 (IND)	2-6	4
		8	Day 8 (IND)	7-9	8
		11	Day 11 (IND)	10-13	11
		15	Day 15 (IND)	14-16	15
		18	Day 18 (IND)	17-20	18
		22	Day 22 (IND)	21- 23	22
		25	Day 25 (IND)	24 to end of IND	25
	OP/MA Weekly x=1, 2, 3, ... , 46	32	Week 1 (OP/MA)	1-8	5
		32 + x*7	Week (x+1) (OP/MA)	2 + x*7 to 8 + x*7	5 + x*7
		361	Week 48 (OP/MA)	331 to end of OP/MA	334
BPIC-SS	IND	1	Baseline (IND)	<= 1	1
		15	Day 15 (IND)	2-21	15
		28	Day 28 (IND)	22 to end of IND	28
		IND final visit	End Point (IND)	2 to end of IND	
	OP/MA Every 4 weeks x=1, 2, 3, … , 10	28	Baseline (OP/MA)	1	1
		53	Week 4 (OP/MA)	2-40	26
		53 + x*28	Week (4 + x*4) (OP/MA)	13 + x*28 to 40 + x*28	26 + x*28
		361	Week 48 (OP/MA)	321 to end of OP/MA	334
		OP/MA final visit	End Point (OP/MA)	2 to end of OP/MA	
	F/U	F/U Day 14 (Last dose +14 days)	F/U Week 2	1 to 21	14
		F/U Day 28 (Last dose +28 days)	F/U Week 4	22 to end of F/U	28
PWC-20	IND	IND final visit	End Point (IND)	2 to end of IND	
	OP/MA	OP/MA final visit	End Point (OP/MA)	1 to end of OP/MA	
	F/U	F/U Day 7 (Last dose +7 days)	F/U Week 1	1 to 10	7
		F/U Day 14 (Last dose +14 days)	F/U Week 2	11 to 21	14
		F/U Day 28 (Last dose +28 days)	F/U Week 4	22 to end of F/U	28
		F/U final visit	End Point (F/U)	1 to end of F/U	

Table 1: Analysis Visits

Parameter	Analysis Phase	Scheduled Day	Time Interval (label on output)	Time Interval[a] (Day)	Target Time Point from start of each phase (Day)
HRUQ	OP/MA Every Two Weeks x=1, 2, 3, … , 20	42	Week 6 (OP/MA)	1-49	42
		56	Week 8 (OP/MA)	50-63	56
		56 + x*14	Week (8 + x*2) (OP/MA)	50 + x*14 to 63 + x*14	56 + x*14
		361	Week 48 (OP/MA)	321 to end of OP/MA	334
		OP/MA final visit	End Point (OP/MA)	2 to end of OP/MA	
	F/U	F/U Day 14 (Last dose +14 days)	F/U Week 2	1 to 21	14
		F/U Day 28 (Last dose +28 days)	F/U Week 4	22 to end of F/U	28
		F/U final visit	End Point (F/U)	1 to end of F/U	

[a] For each phase, the time interval is relative to the first day of that phase.
[b] TSH, HbA1c and Lipid panel (fasting) are only recorded at screening. Summarize the statistics of these variables at screening phase.
[c] During each phase, at 1.5 hours postdose, if the SBP is ≥160 mm Hg and/or the DBP ≥100 mm Hg, assessments should continue every 30 minutes until the blood pressure is <160 mm Hg SBP and <100 mm Hg DBP or until the subject is referred for appropriate medical care if clinically indicated.
[d] If the MOAA/S score is ≤ 3 at any time during the 1.5 hour postdose interval, the MOAA/S will be performed every 5 minutes until a score of 4 is reached (at which point a frequency of every 15 minutes can be resumed until t=+1.5 hours post dose).). If a subject does not have a score of at least 5 at t=+1.5 hours postdose, they should continue to be monitored. For subjects with a score of 4, the assessment should be repeated every 15 minutes. And for subjects with a score of □3, the assessment should be repeated every 5 minutes until the score returns to 5 or the subject is referred for appropriate medical care, if clinically indicated.
[e] If pulse oximetry is <93% at any time during the 1.5 hour postdose interval, pulse oximetry will be performed every 5 minutes until oxygen saturation returns to ≥93% or until the subject is referred for appropriate medical care, if clinically indicated.
[f] If the response is not "Yes" at 1.5 hour postdose, the assessment will be repeated every 15 minutes until a "Yes" response is achieved or until the subject is referred for appropriate medical care if clinically indicated.
[g] This would be labeled as Day 1(IND): Predose for changes from Predose summaries.

2.4. Analysis Sets

Subjects will be classified into the following analysis sets: all enrolled, full analysis sets and follow-up analysis set. Due to Good Clinical Practice (GCP) issues, Dr. [redacted] will not be included in any analysis sets. However, data for this site will be presented in listings. Analyses of change from baseline will include only those subjects who have both baseline and at least 1 post-baseline observation in that phase.

2.4.1. All Enrolled Analysis Set

This analysis set will include all transferred-entry and direct-entry subjects who are not screen failures and received at least one dose of intranasal study medication or 1 dose of oral antidepressant.

2.4.2. Full Analysis Sets

The following analysis sets will be used to summarize efficacy and safety data and are defined for each phase.

Full (IND) analysis set: All subjects who receive at least 1 dose of intranasal study medication or 1 dose of oral antidepressant in the open-label induction phase (for direct-entry and transferred-entry non-responder subjects).

Full (OP/MA) analysis set: All subjects who receive at least 1 dose of intranasal study medication or 1 dose of oral antidepressant in the optimization/maintenance phase.

2.4.3. Follow-up Analysis Set

The Follow-up analysis set includes all subjects who enter the follow-up phase. This analysis set will be used to summarize all efficacy and safety evaluations during the follow-up phase.

2.5. Definition of Subgroups

Descriptive statistics will be provided for the change from baseline in MADRS total score, using the following subgroups.

- Sex
- Race (White, Black, Other)
- Age Group (18-44 years, 45-64 years, 65-74 years, ≥75 years)
- Region (North America, Europe, Asia, Other)
- Country (Brazil, Belgium, France, Poland, Spain, United States, Italy, United Kingdom, Sweden, Mexico, Germany, Bulgaria, Malaysia, Romania, Taiwan, South Africa, Argentina, Austria, Australia, Turkey, Republic of South Korea, Lithuania, Finland)
- Number of Previous Treatment Failures in Current Episode (based on ATRQ)
- Functional Impairment based on Baseline (IND) SDS Total Score: not impaired (0-3), mild (4-11), moderate (12-19), marked (20-26) or extreme (27-30)
- Baseline (IND) MADRS total score (≤/> median)
- Class of antidepressant study medication (SNRI or SSRI)

2.6. Imputation Rules for Missing AE Dates

Treatment-emergent adverse events (AEs):

- For direct-entry and transfer-entry non-responder subjects, treatment-emergent AEs are those events with an onset date/time on or after the start of IND phase study medication, and occurred on or before the end of the optimization/maintenance phase. A conservative approach will be used to handle the missing dates for AEs
- For transferred-entry responder subjects, treatment-emergent AEs are those events with an onset date/time on or after the start of optimization/maintenance phase study medication,

and occurred on or before the end of the optimization/maintenance phase. A conservative approach will be used to handle the missing dates for AEs

Onset Date

AEs for each phase are those events with an onset date/time on or after the start of that particular phase, and occurred before the end of that phase. The rules for estimating incomplete AE onset dates will be as follows:

Direct-entry and transferred-entry non-responder subjects:

If the onset date of an adverse event is missing the day only, it will be set to:

i) First day of the month that the AE occurred, if month/year of the onset of AE is different than the month/year of IND start date

ii) The day of IND start date, if the month/year of the onset of AE is the same as month/year of the IND start date and month/year of the AE resolution date is different

iii) The day of IND start date or day of AE resolution date, whichever is earliest, if month/year of the onset of AE and month/year of the IND start date and month/year of the AE resolution date are the same.

If the onset date of an adverse event is missing both day and month, it will be set to the earliest of:

i) January 1 of the year of onset, as long as this date is after the IND start date

ii) One day after the IND start date, if this date is the same year that the AE occurred.

A completely missing onset date of an adverse event will be set to the IND start date.

Transferred-entry responder subjects:

If the onset date of an adverse event is missing day only, it will be set to:

i) First day of the month that the AE occurred, if month/year of the onset of AE is different than the month/year of OP/MA start date

ii) The day of OP/MA start date, if the month/year of the onset of AE is the same as month/year of the OP/MA start date and month/year of the AE resolution date is different

iii) The day of OP/MA start date or day of AE resolution date, whichever is earliest, if month/year of the onset of AE and month/year of the OP/MA start date and month/year of the AE resolution date are the same.

If the onset date of an adverse event is missing both day and month, it will be set to the earliest of:

i) January 1 of the year of onset, as long as this date is after the OP/MA start date

ii) One day after the OP/MA start date, if this date is the same year that the AE occurred.

A completely missing onset date of an adverse event will be set to the OP/MA start date.

Resolution Date

The missing day of resolution of an adverse event will be set to the last day of the month of resolution.

If the resolution date of an adverse event is missing both day and month, it will be set to December 31 of the year.

A completely missing resolution date of an adverse event that is not recorded as ongoing will be set to the date of withdrawal or study completion.

Direct-entry and transferred-entry non-responder subjects:

If the time of onset is missing, it will be imputed as follows:

(i) 00:00 if the date of onset is after IND start date

(ii) 00:00 if the date is the same as IND start date, but the intranasal study medication in the open-label induction phase was started after the oral antidepressant medication in this phase

(iii) The time of intranasal medication start in the open-label induction phase if the intranasal medication was started on or before the oral antidepressant medication in this phase

Transferred-entry responder subjects:

If the time of onset is missing, it will be imputed as follows:

(i) 00:00 if the date of onset is after OP/MA start date

(ii) 00:00 if the date is the same as OP/MA transferred-entry responder subjects, start date, but the intranasal study medication in the optimization/maintenance phase was started after the oral antidepressant medication in this phase

(iii) The time of intranasal medication start in the optimization/maintenance phase if the intranasal medication was started on or before the oral antidepressant medication in this phase

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

2.7. Imputation Rules for Missing Concomitant Medication Dates

If a partial date is reported, it is assumed the medication (or therapy) was taken in all phases that overlap with the partial date. If both start and end dates are missing but this concomitant medication was taken both prior to the study entry and still ongoing at study end, it is assumed medication was taken in all phases.

The rules for estimating an incomplete concomitant medication start date are as follows:

Direct-entry and transferred-entry non-responder subjects: If the month of the concomitant medication start date is equal to the month of the start of the induction phase, then the estimated start date is the IND start date;

If the month of the concomitant medication start date is greater than the month of the start of the induction phase and earlier than the study end date, then the estimated start date of the concomitant medication is the first day of the month;

If the month of the concomitant medication start date is greater than the month of the study end date, then no imputation will be done;

If the month and year of the concomitant medication start date are known and the IND start date is after the month of the concomitant medication start date, then no imputation will be done;

If either the month or year of the concomitant medication start date is missing, no imputation is to be performed.

Transferred-entry responder subjects:

If the month of the concomitant medication start date is equal to the month of the start of the optimization/maintenance phase, then the estimated start date is the OP/MA start date;

If the month of the concomitant medication start date is greater than the month of the start of the optimization/maintenance phase and earlier than the study end date, then the estimated start date of the concomitant medication is the first day of the month;

If the month of the concomitant medication start date is greater than the month of the study end date, then no imputation will be done;

If the month and year of the concomitant medication start date are known and the OP/MA start date is after the month of the concomitant medication start date, then no imputation will be done;

If either the month or year of the concomitant medication start date is missing, no imputation is to be performed.

3. DATA MONITORING COMMITTEE REVIEW

An IDMC will be established to monitor data on an ongoing basis to ensure the continuing safety of the subjects enrolled in this study. The committee will meet every 6 months to review safety data. After the reviews, the IDMC will make recommendations regarding the continuation of the study.

4. SUBJECT INFORMATION

4.1. Demographics and Baseline Characteristics

Demographic and baseline characteristics (Table 2) and psychiatric history at baseline (Table 3) will be summarized for the All Enrolled analysis set (described in Section 2.4.1). Continuous variables will be summarized using descriptive statistics (N, mean, standard deviation [SD], median, minimum, and maximum). Categorical variables will be summarized using a frequency distribution with the number and percentage of subjects in each category. Baseline (IND) values will be used for these summaries.

Table 2: Demographic Variables and Baseline Characteristics

Continuous Variables:

- Age (years) (informed consent date (for transfer entry subjects use 3004 informed consent date) – date of birth + 1) / 365.25
- Baseline weight (kg)
- Baseline height (cm)
- Baseline BMI (kg/m^2) calculated as Weight (kg)/[Height (m)]^2

Categorical Variables**:**

- Age in years (18-44, 45-64, 65-74, ≥75)
- Sex (male, female)
- Race[a] (White, Black or African American, Asian, American Indian or Alaskan native, Native Hawaiian or other Pacific islander, other)
- Ethnicity (Hispanic or Latino, not Hispanic or Latino)
- Baseline BMI (underweight <18.5kg/m^2; normal: 18.5 to <25 kg/m^2, overweight: 25 kg/m^2 to <30 kg/m^2, obese:30 to <40 kg/m^2; morbidly obese: ≥40 kg/m^2)
- Employment Status
- Hypertension status
- Class of antidepressant (SSRI/SNRI)
- Oral antidepressant
- Country
- Region (North America, Europe, Other)

[a] If multiple race categories are indicated, then Race is recorded as “Multiple”.

Table 3: Psychiatric History at Baseline Variables

Continuous Variables:

- Baseline MADRS total score
- Baseline CGI-S score
- Baseline PHQ-9 total score
- Age (years) when diagnosed with MDD
- Duration of current episode

Categorical Variables:

- Baseline CGI-S category
- Screening C-SSRS category lifetime (no event, suicidal ideation, suicidal behavior)
- Screening C-SSRS category past 6 or 12 months (no event, suicidal ideation [past 6 months], suicidal behavior [past 12 months])
- Antidepressant treatment history (number of medications with non-response taken for at least 2 weeks during the current episode as obtained in the MGH-ATRQ)
- Family history of
  - Depression
  - Anxiety Disorder
  - Bipolar Disorder
  - Schizophrenia
  - Alcohol Abuse
  - Substance Abuse

4.2. Disposition Information

The number of subjects who enrolled from study ESKETINTRD3005 by responder status and direct-entry into ESKETINTRD3004 will be provided.

The following disposition summaries will be provided for each phase separately. These summaries will be provided for each of the full analysis sets described in Section 2.4.

- The number of subjects who entered a specific treatment phase
- The number of subjects who discontinued a specific treatment phase prematurely and their reasons for discontinuation
- The number of subjects who are ongoing in each phase at the time the sponsor terminated the study.

Subjects who terminated the trial and the reasons for ending study participation throughout the study will also be summarized. This will be summarized for All Enrolled analysis sets described in Section 2.4.

4.3. Extent of Exposure

Extent of exposure in terms of total duration of exposure and number of dosing sessions of intranasal study medication will be summarized by phase for the full analysis sets described in Section 2.4.

Extent of exposure in terms of total duration of exposure and number of dosing sessions of intranasal study medication will be summarized by phase for Full (IND) and Full (OP/MA) analysis sets, and across both phases for the All Enrolled described in Section 2.4

The total duration of exposure for the intranasal study drug and for each type of oral antidepressant (AD) during each phase is defined as the time between the first and the last dose of each type of study medication in that specific phase (last day of study medication-first day of study medication +1). If a subject only receives a partial dose it is considered as a day of dosing.

Modal dose for a subject is defined as the most frequently taken dose by a subject during that phase. Mean dose of a subject is calculated as the sum of doses during the phase divided by the total number of days exposed. The final dose is the last non-zero dose received during that phase. The calculation of mean, modal and final dose will exclude days off study drug.

Descriptive statistics (N, mean, SD, median, minimum and maximum) of modal dose, mean dose and final dose will be presented for the induction phase and for the optimization/maintenance phase for intranasal study drug.

Doses of oral AD will be summarized using descriptive statistics of the mean dose (days on drug), final dose and mode dose (days on drug), by each type of oral AD during the induction phase and during the optimization/maintenance phase. The doses will also be presented for direct entry subjects during the induction phase.

At the end of the induction phase and the optimization/maintenance phase, the number and percentage of subjects at each dose (28 mg, 56 mg and 84 mg) will be provided.

A frequency distribution of subjects with 6 months (≥180 days) of exposure and 12 months (≥350 days) of exposure will be provided.

During the optimization/maintenance phase, the number and percentage of subjects at each dosing frequency (weekly or every other week) and the number of subjects who changed their frequency (weekly to every other week and every other week to weekly) will be summarized every 4 weeks starting at Week 8.

4.4. Protocol Deviations

Deviations that occurred during the study will be tabulated for the All Enrolled analysis set. Major deviations will be tabulated for the following categories: subject not withdrawn as per protocol, selection criteria not met, excluded concomitant treatment, treatment deviation, non-compliance, regulatory requirement. More categories may be included depending on the nature of the protocol deviation.

4.5. Prior and Concomitant Medications

The number and percent of subjects receiving prior antidepressant medications/therapies will be summarized for the direct-entry subjects using Full (IND) analysis set described in Section 2.4.

The number and percent of subjects who receive concomitant therapies will be summarized by phase using the generic term of the medication for the Full (IND, OP/MA) analysis sets and the Follow-up analysis set described in Section 2.4.

5. EFFICACY

Efficacy measures will be summarized descriptively at each scheduled visit for each phase, using both last observation carried forward and observed data.

5.1.1. Data Handling Rules

For the efficacy scales MADRS, CGI-S, PHQ-9, GAD-7, and SDS, both observed case and last observation carried forward (LOCF) values will be determined for the induction and optimization/maintenance phases. The last post baseline observation during the phase will be carried forward as the "End Point" for that phase. Besides the observed cases and the end point assessment, the LOCF values will be created for intermediate postbaseline time points as well. These imputed time points will be labeled 'DAY X LOCF' or 'WEEK X LOCF'.

5.1.2. Imputation Methods for Missing Items

Imputation of missing individual item scores will apply to MADRS and is described in Section 5.2.1. For all other scales where multiple items are summed to create a total, if any item of the scale is missing at a visit, the total score for that scale at that visit will be considered missing.

5.2. Efficacy Endpoint(s)

Efficacy analyses during the induction phase and the optimization/maintenance phase will be provided for the Full analysis sets defined in Section 2.4.2 and for the follow-up phase will be provided for the Follow-up analysis set defined in Section 2.4.3.

5.2.1. MADRS

5.2.1.1. Definition

The Montgomery-Asberg Depression Rating Scale (MADRS)[6] is a clinician-rated scale designed to measure depression severity and to detect changes due to antidepressant treatment. The scale consists of 10 items, each of which is scored from 0 (item is not present or is normal) to 6 (severe or continuous presence of the symptoms), for a total possible score of 60. Higher scores represent a more severe condition. The MADRS evaluates apparent sadness, reported sadness, inner tension, sleep, appetite, concentration, lassitude, interest level, pessimistic thoughts, and suicidal thoughts. The test exhibits high inter-rater reliability.

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 non-missing items multiplied by the ratio of the maximum number of items (i.e., 10) to the number of non-missing items.

A subject is defined a responder at a given time point if the percent improvement in MADRS total score is ≥50%.

A subject is defined as a remitter at a given time point if the MADRS total score of ≤12 at that time point.

5.2.1.2. Analysis Methods

Descriptive statistics of the total score and change from baseline (of the respective phase) will be provided for each visit during all phases, including the follow-up phase. Graphical presentations will be provided. In addition, the proportion of subjects who responded and remitted based on the MADRS total score will be provided over time for each phase. Summaries of both observed and LOCF data will be presented.

5.2.2. PHQ-9

5.2.2.1. Definition

The PHQ-9 is a 9-item, self-report scale assessing depressive symptoms. 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 0-27. A higher score indicates greater severity of depression. The recall period is 2 weeks. The scale scores each of the nine symptom domains of the DSM MDD criteria and it has been used both as a screening tool and a measure of response to treatment for depression. The severity of the PHQ-9 is categorized as follows: None-minimal (0-4), Mild (5-9), Moderate (10-14), Moderately Severe (15-19) and Severe (20-27).

A subject is defined a responder at a given time point if the percent improvement from Baseline (IND) in PHQ-9 total score is ≥50%.

A subject is defined as a remitter at a given time point if the PHQ-9 total score is ≤5 at that time point.

5.2.2.2. Analysis Methods

Descriptive statistics of the total score and change from baseline (of the respective phase) will be provided for each visit during during all phases, including the follow-up phase Graphical presentations will be provided. Summaries of both observed and LOCF data will be presented.

In addition, a frequency distribution by severity will be provided for each assessment visit for all phases, including the follow-up phase. In addition, the proportion of subjects who responded and remitted based on the PHQ-9 total score will be provided over time for each phase. Summaries of both observed and LOCF data will be presented.

5.2.3. CGI-S

5.2.3.1. Definition

The Clinical Global Impression of Severity (CGI-S)[4] provides an overall clinician-determined summary measure of the severity of the subject's illness that takes into account all available information, including knowledge of the subject's history, psychosocial circumstances, symptoms, behavior, and the impact of the symptoms on the subject's ability to function. The CGI-S evaluates the severity of psychopathology on a scale of 0 to 7. Considering total clinical

experience, a subject is assessed on severity of mental illness at the time of rating according to: 0=not assessed; 1=normal (not at all ill); 2=borderline mentally ill; 3=mildly ill; 4=moderately ill; 5=markedly ill; 6=severely ill; 7=among the most extremely ill patients. The CGI-S permits a global evaluation of the subject's condition at a given time.

5.2.3.2. Analysis Methods

Descriptive statistics of the total score and change from baseline (of the respective phase) will be provided for each visit during all phases, including the follow-up phase. Summaries of both observed and LOCF data will be presented.

In addition, a frequency distribution by severity will be provided for each visit for all phases, including the follow-up phase. Summaries of both observed and LOCF data will be presented.

5.2.4. GAD-7

5.2.4.1. Definition

The GAD-7 (Generalized Anxiety Disorder - 7 Items)[9] is a brief and validated 7-item self-report assessment of overall anxiety. Subjects respond to each item using a 4 point scale with response categories of 0=not at all, 1=several days, 2=more than half the days, and 3=nearly every day. Item responses are summed to yield a total score with a range of 0 to 21, where higher scores indicate more anxiety. The recall period is 2 weeks. The severity of the GAD-7 is categorized as follows: None (0-4), Mild (5-9), Moderate (10-14) and Severe (15 -21).

5.2.4.2. Analysis Methods

Descriptive statistics of the actual values and change from baseline (of the respective phase) will be provided for each visit during all phases, including the follow-up phase. Graphical presentations will be provided. Summaries of both observed case data and LOCF will be presented.

In addition, a frequency distribution of GAD-7 severity categories will be provided for each assessment visit for all phases, including follow-up phase. Summaries of both observed and LOCF data will be presented.

5.2.5. Sheehan Disability Scale (SDS)

5.2.5.1. Definition

The SDS is a subject-reported outcome measure and is a 5-item questionnaire which has been widely used and accepted for assessment of functional impairment and associated disability. The first three items assess disruption of (1) work/school, (2) social life, and (3) family life/home responsibilities using a 0-10 rating scale. The score for the first three items are summed to create a total score of 0-30 where a higher score indicates greater impairment. It also has one item on days lost from school or work and one item on days when underproductive. The recall period for this study is 7 days. Scores ≤4 for each item and ≤12 for the total score are considered response. Scores ≤2 for each item and ≤6 for the total score are considered remission. If any of the first

three items are missing, the total score will be set to missing as well as response and remission status.

5.2.5.2. Analysis Methods

Descriptive statistics of the actual values and change from baseline (of the respective phase) will be provided for each visit during all phases, including the follow-up phase. Graphical presentations will be provided. Summaries of both observed and LOCF data will be presented.

The total score as well as the individual item scores will be summarized. In addition, the proportion of subjects who achieve response and remission will be summarized at each time point.

5.2.6. EuroQol Group; 5 Dimension; 5 Level (EQ-5D-5L)

5.2.6.1. Definition

The EQ-5D-5L (EuroQol Group - 5 Dimension - 5 Level) [2,3] is a standardized 2-part instrument for use as a measure of health outcome, primarily designed for self-completion by respondents. It essentially consists of the EQ-5D-5L descriptive system and the EQ visual analogue scale (EQ VAS). The EQ-5D-5L descriptive system comprises the following 5 dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each of the 5 dimensions is divided into 5 levels of perceived problems (Level 1 indicating no problem, Level 2 indicating slight problems, Level 3 indicating moderate problems, Level 4 indicating severe problems, and Level 5 indicating extreme problems).

The subject selects an answer for each of the 5 dimensions considering the response that best matches his or her health "today." The descriptive system can be represented as a health state. The EQ VAS self-rating records the respondent's own assessment of his or her overall health status at the time of completion, on a scale of 0 (the worst health you can imagine) to 100 (the best health you can imagine).

The time taken to complete the questionnaire varies with age, health status, and setting but is likely to be around 1 minute.

Individual scores from the 5 dimensions will be used to obtain a weighted health status index as shown below:

(i) Scores from each dimension will be combined to obtain a 5L profile score or health state: eg, a score of 1 for each dimension will give a 5L profile score of 11111. Dimension scores will be combined in the following order: Mobility, Self-Care, Usual Activities, Pain/Discomfort, Anxiety/Depression

(ii) The value set of the Health Status Index for various values of 5L profile scores is published for Canada in the following website: https://www.ncbi.nlm.nih.gov/pubmed/26492214

(iii) The Canadian value set will be used to get the HSI values for all the countries participating in the study.

In addition, a sum score will be derived as follows: The scores of the five dimensions (values 1-5) will be added (sums between 5 and 25). From this score, subtract 5 (range 0-20) and multiply by 5 (range 0-100).

5.2.6.2. Analysis Methods

Descriptive statistics of actual values and changes from baseline for the weighted EQ-5D health status index, the EQ-VAS, and the sum score will be summarized for each visit for all phases including the follow-up phase. In addition, individual dimension responses using a frequency distribution will also be summarized for each visit for all phases, including the follow-up phase. Summaries of observed data will be presented.

6. SAFETY

Safety data for the open-label induction phase, the optimization/maintenance phase and follow-up phase will be summarized separately for each phase. Summaries will be based on Full analysis sets and the Follow-up analysis set described in Section 2.4. In addition, adverse event and cognition data, and most severe postbaseline C-SSRS, treatment-emergent abnormal vital signs, ECG, laboratory, and MOAS/S data will be summarized for the entire treatment period combining the induction and optimization/maintenance phase. All enrolled analysis set will be used for summarizing the combined treatment phases.

6.1. Adverse Events

The Medical Dictionary for Regulatory Activities (MedDRA) (version 18.1 or above) will be used to classify AEs by system organ class and preferred term. Treatment-emergent adverse events (TEAEs) that occurred in each study phase will be summarized by system organ class and preferred term.

The number (%) of subjects with TEAEs, serious TEAEs (SAEs), and TEAEs that led to study drug discontinuation will be summarized by system organ class and preferred term. Data listings will be generated for deaths, other SAEs, and discontinuations due to AEs.

For transferred-entry subjects, the number (%) of adverse events ongoing at the time of entry into the ESKETINTRD3004 study will be listed separately.

A TEAE is an event that is new in onset or increased in severity following treatment initiation. 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.

TEAEs are defined as follows for each study phase:

- TEAEs in the induction phase:
  a. If AE onset time is not missing:
     i. If subjects continue to OP/MA phase: IND phase start date/time <= AE onset date and time < IND phase end date

ii. If subjects discontinue in the IND phase: IND phase start date/time <= AE onset date and time <= IND phase end date

b. If AE onset time is missing:

i. If subjects continue to OP/MA phase: IND phase start date <= AE onset date < IND phase end date

ii. If subjects discontinue in the IND phase: IND phase start date <= AE onset date <= IND phase end date

- TEAEs in the optimization/maintenance phase:
  a. If AE onset time is not missing: OP/MA phase start date ≤ AE onset date ≤ OP/MA phase end date
  b. If AE onset time is missing: OP/MA phase start date ≤ AE onset date ≤ OP/MA phase end date
- AEs in follow-up phase: F/U start date ≤ AE onset date ≤ F/U end date
- For the AEs that have both day and month missing, treatment emergent flag is be assigned based on the rules presented in Section 2.6.

In addition, TEAEs will be summarized by severity and relationship to study medication using the preferred term. For the summaries of AEs by severity/relationship to study medication, the observation with the most severe occurrence/closest relationship to study medication will be chosen if there is more than one incident of an adverse event reported during the analysis phase by the subject. AE duration for transient dizziness/vertigo and anxiety will also be summarized.

Adverse Events of Special Interest

Clinically relevant TEAEs of special interest will be examined separately grouped in the following categories:

- drug abuse, dependence and withdrawal (Aggression, Confusional state, Decreased activity, Dependence, Disorientation, Dissociation, Dissociative disorder, Dizziness, Drug use disorder, Drug abuse, Drug abuser, Drug dependence, Drug detoxification, Drug diversion, Drug rehabilitation, Drug tolerance, Drug tolerance increased, Drug withdrawal convulsions, Drug withdrawal headache, Drug withdrawal syndrome, Euphoric mood, Feeling abnormal, Feeling drunk, Feeling of relaxation, Hallucination, Hallucination, auditory, Hallucination, gustatory, Hallucination, olfactory, Hallucination, synaesthetic, Hallucination, tactile, Hallucination, visual, Hallucinations, mixed, Inappropriate affect, Mental impairment, Product tampering, Psychomotor hyperactivity, Psychotic disorder, Rebound effect, Somatic hallucination, Somnolence, Substance abuser, Substance dependence, Substance use, Substance use disorder, Substance-induced mood disorder, Substance-induced psychotic disorder, Thinking abnormal, Withdrawal arrhythmia, Withdrawal syndrome);
- increased blood pressure (Blood pressure increased, Blood pressure diastolic increased, Blood pressure systolic increased, Hypertensive crisis, Hypertensive emergency, Hypertension);
- increased heart rate (Heart rate increased, Tachycardia);

- transient dizziness/vertigo (Dizziness, Dizziness exertional, Dizziness postural, Procedural dizziness, Vertigo, Vertigo labyrinthine, Vertigo positional, Vertigo CNS origin);
- impaired cognition (Cognitive disorder);
- cystitis (Allergic cystitis, Chemical cystitis, Cystitis, Cystitis erosive, Cystitis haemorrhagic, Cystitis interstitial, Cystitis noninfective, Cystitis ulcerative, Cystitis-like symptom);
- anxiety (Anticipatory anxiety, Anxiety, Anxiety disorder).

The number and percentage of subjects taking concomitant medication for dissociation events (preferred term of Dissociation) at any time during each treatment phase will be provided.

Summary statistics for the duration of all episodes of TEAEs associated with discharge readiness (Dissociation, Dizziness, Feeling abnormal, Feeling drunk, Nausea, Somnolence, Vertigo, and Vomiting) with an onset on the day of intranasal study drug administration is summarized for each treatment phase. These summaries are presented for each dosing session during both the induction phase and for the optimization/maintenance phase. In addition, the number of occurrences of TEAEs associated with discharge readiness and the number of occurrences of TEAEs associated with substance dependence is presented by dosing session for each treatment phase.

6.2. Clinical Laboratory

Descriptive statistics (N, mean, median, minimum, and maximum and range) for values and changes from baseline will be provided for clinical laboratory tests (hematology, chemistry and urinalysis) at each scheduled time point for all phases, including the follow-up phase. Baseline (IND) will be used for change summaries and to determine abnormal values during all treatment phases and during the follow-up phase.

Clinical laboratory tests that meet the criteria for markedly abnormal will be listed by subject for each phase. The incidence of treatment emergent markedly abnormal (TEMA) laboratory values that occurred at any time during each treatment phase will be presented. Clinical laboratory test values will be considered "TEMA using the criteria defined by the Sponsor (Janssen Research & Development, LLC)" listed in Attachment 1. The identification of TEMA laboratory values is based on the postbaseline value being out of range while the baseline value (Baseline (IND)) is either missing or within the range given in Attachment 1. 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 postbaseline value being below the lower limit with the baseline value being above the upper limit. Baseline (IND) will be used.

In addition, incidence of subjects with ALT values >3*upper normal limit (ULN) will be presented for each study phase. Additionally, incidence of hepatic toxicity (Hy's Law)[10] defined as ALT values >3*ULN and total bilirubin values >2*ULN will be presented for each study phase. Similar to the markedly abnormal analysis, only subjects with Baseline (IND) ALT values ≤3*ULN (and baseline (IND) total bilirubin values ≤2*ULN for hepatic toxicity) (or if the baseline (IND) value is missing) will be eligible for these analyses.

6.3. Vital Signs, Weight, and BMI

Descriptive statistics for values and changes from Baseline (IND) at each scheduled time-point during the open-label induction phase, the optimization/maintenance phase, the follow-up phase will be presented for temperature, systolic blood pressure, diastolic blood pressure, pulse rate, respiratory rate, oxygen saturation, weight, and BMI. Baseline (IND) will be used for change summaries and to determine abnormal values during all treatment phases and during the follow-up phase. Graphical presentations will be provided.

Vital sign mean and mean changes from Baseline (IND) at each scheduled time point during each phase will also be provided by sex and dose level and by age group (<65 years; ≥65 years) and dose level.

In addition, descriptive statistics of pulse rate and blood pressure (systolic and diastolic) values and changes and percent changes from predose will be provided for each intranasal dosing day.

The proportion of subjects who have a treatment-emergent abnormality, as defined in Table 4 below, during each treatment phase will be presented. A listing of subjects meeting any of the criteria will also be provided.

Table 4: Treatment-Emergent Abnormality Categories for Vital Signs

	Post-baseline value outside of normal limit if:	
Vital Parameter	**Abnormally low**	**Abnormally high**
Pulse (bpm)	A decrease from baseline of ≥15 to a value ≤50	An increase from baseline of ≥15 to a value ≥100
Systolic BP (mmHg)	A decrease from baseline of ≥20 to a value ≤90	An increase from baseline of ≥20 to a value ≥180
Diastolic BP (mmHg)	A decrease from baseline of ≥15 to a value ≤50	An increase from baseline of ≥15 to a value ≥105

BP = blood pressure

The proportion of subjects who experienced treatment-emergent acute hypertension (systolic BP≥180 or diastolic BP≥110) at any time during each treatment phase will be summarized by hypertension status.

Mean (+/-SE) values for systolic BP, diastolic BP and heart rate will be summarized by hypertension status and presented graphically for each treatment phase. In addition, for subjects with hypertension who receive antihypertensive medication, the same tables and graphs will be summarized by medication type (beta-blockers, all other agents)

A listing of subjects with oxygen saturation less than 93% will be provided.

6.4. Electrocardiogram

ECG variables that will be analyzed include heart rate, RR, PR interval, QRS interval, QT interval and QTc intervals. The corrected QT (QTc) intervals will include QTcB (Bazett) and QTcF (Fridericia).

Baseline ECG is defined as the average of all ECG results collected up to and including the day of first dose of study medication (either intranasal or oral AD) for direct entry and transferred-entry responders. Baseline is the last observation prior to or on the start date of the open-label induction phase for transferred-entry nonresponders.

The maximum post-baseline value during each treatment phase (open-label induction phase, and optimization/maintenance phase) will be presented separately and will be computed for each ECG parameter using data from both scheduled and unscheduled visits.

Summary tables for observed values and changes from baseline will be presented at each scheduled time point during both treatment phases (open-label induction phase and optimization/maintenance phase), and during the follow-up phase.

The frequency of treatment-emergent abnormalities will be tabulated and presented for all treatment phases. The identification of treatment-emergent abnormal ECG values is based on the post-baseline value (a value occurring after the start of the phase) being out of range while the baseline value is either missing or within the limits given in Table 5. If post-baseline ECG results are above the upper limits (abnormally high) and the baseline value is below the lower limits (abnormally low), then the post-baseline abnormality will also be considered treatment-emergent. The same applies to the post-baseline value being below the lower limits (abnormally low) with the baseline value being above the upper limits (abnormally high). Abnormal ranges for the HR, PR, QRS and QT intervals are given in Table 5.

Table 5: Limits for HR, PR, QRS and QT Interval Abnormality

ECG parameter	Abnormally Low	Abnormally High
HR (bpm)	≤ 50	≥ 100
PR interval (msec)	--	≥ 210
QRS interval (msec)	≤ 50	≥ 120
QT interval (msec)	≤ 200	≥ 500

Based on the maximum QTc value for each subject during a given phase (separate for each QTc correction QTcB and QTcF) the incidence of abnormal QTc values and changes from baseline will be summarized. Criteria for abnormal corrected QT intervals and changes from baseline are given in Table 6 and are derived from the ICH E14 Guidance[5] (the same criteria apply to all QT corrections).

Table 6: Criteria for Abnormal QTc Values and Changes From Baseline

Parameter	Classification	Criteria
Clinically Significant QTc Value	No	≤500
	Yes	>500
QTc change from baseline[a]	No concern	≤30
	Concern	>30 – 60
	Clear concern	>60
QTc value	Normal	≤450
	>450 – 480	>450 - ≤480
	>480 – 500	>480 – ≤500

	>500	>500

These criteria are based on ICH E14 Guideline
[a] Baseline is defined as the average pre-dose.

The proportion of subjects with treatment emergent abnormalities will be presented for both treatment phases. A listing of subjects with abnormalities will also be provided.

6.5. Nasal Examination

Targeted nasal examinations (including the upper respiratory tract/throat) will be conducted by a qualified healthcare practitioner. The objective of the examination at Screening is to rule out any subjects with anatomical or medical conditions that may impede drug delivery or absorption.

Subsequent examinations will consist of a visual inspection of the nostrils, nasal mucosa, and throat for nasal erythema, rhinorrhea, rhinitis, capillary/blood vessel disruption and epistaxis and graded as follows: absent, mild, moderate, or severe.

Changes in findings from baseline for each examination (including the upper respiratory tract/throat) will be listed for each treatment phase. Baseline (IND) will be used as baseline for change summaries.

6.5.1. Nasal Symptom Questionnaire

Subjects will complete a nasal symptom questionnaire on every dosing day at predose and again at 1 hour postdose. The questionnaire was developed to assess nasal tolerability following intranasal administration of study drug. Subjects will rate nasal symptoms as none, mild, moderate, or severe for the following items: stuffy nose, blocked nose, runny nose, itching nose, crusting discharge in or on nose, dryness of nose, burning sensation in the nose, discomfort of nose, bleeding from the nose, postnasal drip, cough, sore throat, taste disturbance and sneezing.

Frequency distributions will be provided for each of the items at each time point during each treatment phase. Shift from pre-dose to post-dose questionnaires during each time point throughout the study will be provided to see if there is any change after repeated administrations of study drug. Frequency of subjects who report moderate/severe symptoms at any postdose timepoint will be presented for each treatment phase. In addition, a listing of severe symptoms will also be presented.

6.6. Other Safety Parameters

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

The Columbia Suicide Severity Rating Scale (C-SSRS)[7] is a low-burden measure of the spectrum of suicidal ideation and behavior that was developed in the National Institute of Mental Health Treatment of Adolescent Suicide Attempters Study to assess severity and track suicidal events through any treatment. 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. Missing scores will not be imputed.

The summaries of the C-SSRS outcomes will be based on the Full analysis set for subjects who have at least 1 post-baseline C-SSRS measurement and a pre-treatment C-SSRS assessment.

A frequency distribution at each scheduled time point will be provided. Shifts from the Baseline (IND) value to the most severe/maximum score during each phase, will be summarized.

The maximum score assigned for each subject will also be summarized into one of three broad categories: No suicidal ideation or behavior (0), Suicidal ideation (1-5), Suicidal behavior (6-10). Shifts from the Baseline (IND) value to the maximum category during each phase will be summarized.

6.6.2. Clinician Administered Dissociative States Scale (CADSS)

The CADSS (Clinician Administered Dissociative States Scale) [1] is an instrument for the measurement of present-state dissociative symptoms, and is administered to assess treatment-emergent dissociative symptoms. The CADSS comprises 23 subjective items and participant's responses are coded on a 5-point scale (0 = "Not at all", 1 = "Mild", 2 = "Moderate", 3 = 'Severe" and 4 = "Extreme"). The CADSS is divided into 3 components using the scoring method shown in Table 7.

Table 7: CADSS Scoring

Component	Questions	Range
Depersonalization	Sum of 3, 4, 5, 6, 7, 20, 23	0-28
Derealization	Sum of 1, 2, 8, 9, 10, 11, 12, 13, 16, 17, 18, 19, 21	0-52
Amnesia	Sum of 14, 15, 22	0-12
Total Score	Sum of 1 through 23	0-92

For the total score and each component, a higher score represents a more severe condition. If any of the response is missing, then the total score is set to missing. The CADSS is measured prior to each dose, at 40 minutes, and at 1.5 hours postdose.

Descriptive statistics (N, mean, SD, median, minimum, and maximum) of the total scores and component scores at each time point and visit, changes from pre-dose and proportion of subjects with an increase in CADSS total score from the pre-dose value at any time during the study will be summarized. Mean change in CADSS from pre-dose value will be presented graphically for each intranasal dose day.

6.6.3. Brief Psychiatric Rating Scale (BPRS+)

The Brief Psychiatric Rating Scale (BPRS+) is an 18 item rating scale which is used to assess potential treatment-emergent psychotic symptoms. The BPRS assesses a range of psychotic and affective symptoms rated from both observation of the subject and the subject's own report. Only the four-item positive symptom subscale (BPRS+) will be used in the study to assess treatment-emergent psychotic symptoms. The BPRS+ consists of: suspiciousness, hallucinations, unusual thought content and conceptual disorganization. Each symptom is rated on a scale of 0 to 6 as follows: 0: not present, not evident or absent; 1: very mild; 2: mild; 3: moderate; 4: moderate severe; 5: severe; or 6: extreme. A total score will be derived by summing the individual items, with a range of 0 to 24 with a higher score representing a more severe condition.

The BPRS+ is measured prior to each dose, at 40 minutes, and at 1.5 hours post dose during each of the dosing phase.

Descriptive statistics (N, mean, SD, median, minimum, and maximum) of the total scores at each time point, change from the pre-dose time point within each visit, and the proportion of subjects with an increase in BPRS+ from the pre-dose value at any time during the study will be provided for each treatment phase. The proportion of subjects with a total score of 3 or more at any time during the study will also be provided. Mean (SE) BPRS+ values from pre-dose value will be presented graphically for each intranasal dosing day.

6.6.4. Modified Observer's Assessment of Alertness/Sedation (MOAA/S)

The MOAA/S (Modified Observer's Assessment of Alertness/Sedation) will be used to measure treatment-emergent sedation with correlation to levels of sedation defined by the American Society of Anesthesiologists (ASA) continuum. The MOAA/S scores range from 0 [No response to painful stimulus; corresponds to ASA continuum for general anesthesia] to 5 [Readily

responds to name spoken in normal tone (awake); corresponds to ASA continuum for minimal sedation].

The MOAA/S is measured on each dosing day every 15 minutes from pre-dose to 1.5 hours post dose or longer, if necessary, until the subject has a score of 5.

- If the score is ≤3 at any time during the 1.5 hour postdose interval, the MOAA/S will be performed every 5 minutes until a score of 4 is reached (at which point a frequency of every 15 minutes can be resumed until t=+1.5 hours post dose).
- If a subject does not have a score of 5 at t=+1.5 hours postdose, they should continue to be monitored. For subjects with a score of 4, the assessment should be repeated every 15 minutes. And for subjects with a score of ≤3, the assessment should be repeated every 5 minutes until the score returns to 5 or the subject is referred for appropriate medical care, if clinically indicated.

Descriptive statistics of the MOAA/S score and changes from predose will be summarized at each scheduled time point. Additionally, the proportion of subjects experiencing sedation (score less than or equal to 3) will be summarized at each scheduled time point.

MOAA/S score will be presented graphically for each intranasal dosing day.

6.6.5. Clinical Global Assessment of Discharge Readiness (CGADR)

The Clinical Global Assessment of Discharge Readiness (CGADR) will be used to measure a subject's current clinical status and is the clinician's assessment of the readiness to be discharged from the study site.

The clinician will answer "Yes" or "No" to the question "Is the subject considered ready to be discharged based on their overall clinical status (e.g., sedation, blood pressure, and other adverse events)?"

On each intranasal dosing day, the CGADR will be performed at 1 hour and 1.5 hours postdose, repeated every 15 minutes if necessary until the response is 'Yes'. A subject should not be discharged prior to the 1.5-hour time point.

The proportion of subjects with a response of 'No' at each time point will be presented during each treatment phase.

6.6.6. Physician Withdrawal Checklist (PWC-20)

The PWC-20 is a 20-item simple and accurate method to assess potential withdrawal symptoms following cessation of intranasal esketamine treatment. The PWC-20 will be performed for all subjects on Day 25 to establish a baseline prior to discontinuation of esketamine treatment – although only relevant for those subjects not continuing to the next phase. For those subjects who proceed to the optimization/maintenance phase, the PWC-20 is conducted at the End of Study Visit. If subjects withdraw early from the study during any phase, the PWC-20 will be conducted at the Early Withdrawal Visit.

The proportion of subjects with withdrawal symptoms at the end of each treatment phase (IND, OP/MA) and during the follow-up phase will be presented. In addition, symptoms at follow-up will be compared to the last assessment in the relevant treatment phase (IND, OP/MA) for subjects who discontinue during that treatment phase and will be summarized using the following categories: new or worsened symptoms, symptoms present and unchanged, no symptoms, and improved.

6.6.7. Computerized Cognitive Battery and Hopkins Verbal Learning Test-Revised (HVLT-R)

The effect of intranasal esketamine on cognition will be assessed using the computerized cognitive battery and Hopkins Verbal Learning Test-Revised (HVLT-R).

The computerized cognitive battery provides assessment of multiple cognitive domains, including attention, visual learning and memory, and executive function. The tests use culture-neutral stimuli, enabling use in multilingual/multicultural settings. The HVLT-R is a measure of verbal learning and memory and is a 12-item word list recall test. The total number of correct responses are captured for 4 trials as well as the number of true-positive responses and false-positive errors in Trial 4 (Delayed Recall). The Total Recall will be derived as the sum of trials 1, 2 and 3. Retention % will be derived as (the number of correct responses in Trial 4)/(higher score of Trials 2 and 3) X 100. Recognition Discrimination Index will be derived as the total number of true-positives – the total number of false-positives.

The computerized cognitive battery and HVLT-R will be assessed at Day 1 predose, Day 28/ET of the open-label induction phase, Week 16 of the optimization phase, and every 12 weeks during the maintenance phase. It will also be assessed again at week 4 post-treatment during the follow-up phase.

See Attachment 2 for details of this analysis.

7. HEALTH ECONOMICS

7.1. Healthcare Resource Use Questionnaire (HRUQ)

Medical resource utilization data, associated with medical encounters, will be collected using the Healthcare Resource Use Questionnaire (HRUQ) during the optimization/maintenance phase and follow-up phase bi-weekly.

The number and percentage of subjects who visited at least one healthcare professional or had a hospital emergency room visit because of their depression in the 2 weeks preceding each time point will be summarized for each type of healthcare professional. The total number of visits to each type of healthcare professional, total number of visits to any healthcare professional, and the total number of hospital emergency room visits related to depression in the 2 weeks preceding each time point will be summarized at each time point with descriptive statistics.

REFERENCES

1. Bremner JD, Krystal JH, Putnam FW, Southwick SM, Marmar C, Charney DS, Mazure CM. Measurement of dissociative states with the Clinician-Administered Dissociative States Scale (CADSS). J Trauma Stress. 1998;11(1):125-136.
2. EuroQol Group. About EQ-5D. http://www.euroqol.org/about-eq-5d.html (accessed 18 June 2014).
3. EuroQol Group. EQ-5D-5L User Guide: basic information on how to use the EQ-5D-5L instrument. Version 2.0; October 2013. http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/UserGuide_EQ-5D-5L_v2.0_October_2013.pdf (accessed 6 August 2014)
4. Guy W. ECDEU Assessment Manual for Psychopharmacology. Rockville, MD: U.S. Department of Health, Education, and Welfare; 1976.
5. International Conference on Harmonization (ICH) Tripartite Guideline E14. London: s.n., 2005. pp. The clinical evaluation of QT/QTc interval prolongation and proarrhythmic for non-antiarrhythmic drugs.
6. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. Br J Psychiatry. 1979;134:382-389.
7. Posner K, Oquendo MA, Gould M, Stanley B, Davies M. Columbia Classification Algorithm of Suicide Assessment (C-CASA): classification of suicidal events in the FDA's pediatric suicidal risk analysis of antidepressants. Am J Psychiatry. 2007;164:1035-1043.
8. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. JAMA. 1999; 282:1737-1744.
9. Spitzer RL, Kroenke K, Williams JB, Lowe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med. 2006;166(10):1092-1097.
10. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER). [2009] Guidance for industry drug induced liver injury: premarketing clinical evaluation. Available: http://www.fda.gov/downloads/Drugs/.../Guidances/UCM174090.pdf.

ATTACHMENTS

Attachment 1: Criteria of Markedly Abnormal Laboratory Values

	Markedly Abnormal Limits	
Laboratory Parameter	Low	High
Albumin [g/L]	24	60
Alkaline phosphatase [U/L]	N/A	250
Alanine transaminase (SGPT) [U/L]	N/A	200
Alanine transaminase (SGPT) [U/L]	N/A	>3X ULN
Aspartate transaminase (SGOT) [U/L]	N/A	250
Bicarbonate [mmol/L]	15.1	34.9
Blood urea nitrogen [mmol/L]	N/A	17.9
Calcium [mmol/L]	1.5	3
Chloride [mmol/L]	94	112
Creatine kinase (U/L)	N/A	990
Creatinine [µmol/L]	N/A	265.2
Gamma glutamyl transferase [U/L]	N/A	300
Glucose [mmol/L]	2.2	16.7
Phosphate [mmol/L]	0.7	2.6
Potassium [mmol/L]	3.0	5.8
Sodium [mmol/L]	125	155
Bilirubin, total [µmol/L]	N/A	51.3
Protein, total [g/L]	50	N/A
Urine pH	N/A	8.0
Hematocrit [fraction] - female	0.28	0.5
- male	0.24	0.55
Hemoglobin [g/L]	80	190
Neutrophils, segmented [%]	30	90
Monocytes [%]	N/A	20
Eosinophils [%]	N/A	10
Basophils [%]	N/A	6
Lymphocytes [%]	10	60
Platelet count [$x10^9$/L]	100	600
Erythrocytes (RBC) [$x10^{12}$/L] -- female	3.0	5.5
-- male	3.0	6.4
Leukocyted (WBC) [$x10^9$/L]	2.5	15.0
Hy's Law criteria:		
Alanine transaminase (SGPT) [U/L]		>3X ULN
AND		
Bilirubin, total [µmol/L]		>2X ULN

Note: The same limits apply to both males and females unless gender is indicated;
N/A = Not applicable.

Attachment 2: Statistical Analysis Plan for COGSTATE

STATISTICAL ANALYSIS PLAN

RANDOMIZED, DOUBLE-BLIND, MULTICENTER, ACTIVE-CONTROLLED STUDIES TO EVALUATE THE EFFICACY, SAFETY, AND TOLERABILITY OF FIXED DOSES OF INTRANASAL ESKETAMINE PLUS AN ORAL ANTIDEPRESSANT IN ADULT SUBJECTS WITH TREATMENT-RESISTANT DEPRESSION

AND

AN OPEN-LABEL, LONG-TERM, SAFETY AND EFFICACY STUDY OF INTRANASAL ESKETAMINE IN TREATMENT-RESISTANT DEPRESSION

PROTOCOL ESKETINTRD3001/2/3/4/5; PHASE 3

Prepared for: Janssen Research & Development, LLC

Prepared by: Cogstate Biostatistics Group
Level 8, 195 Church Street
New Haven, CT, USA, 06510

Version: V2
Date: 30-Mar-2016
Protocol Version: Approved Date: 17 Feb 2016

1

1 NOTE

Cogstate has prepared a Statistical Analysis Plan (SAP) for the Sponsor to review and sign-off for all ESKETINTRD studies. Analyses will be provided after this document has been finalized and officially signed. In this SAP, anything in italics is taken directly from the protocol. For more details, please refer to the study protocols and SAPs.

Notice that this SAP will be used for all the five ESKETINTRD studies (3001, 3002, 3003, 3004 and 3005).

2 SIGNATURE PAGE FOR SAP APPROVAL

The following signatures indicate the approval of the statistical analysis plan for ESKETINTRD studies.

Approved by:

Name (print): [redacted]

Position:

Signature:

Date (ddmmmyyyy): 30MAR2016

Approved by:

Name (print): Jaskaran Singh

Position: [redacted]

Signature: Digitally signed by 3D900151-EFFD-4D56-9B44-D29916A7A91E
DN: cn=3D900151-EFFD-4D56-9B44-D29916A7A91E, o=Johnson & Johnson
Date: 2016.03.31 09:05:21 -05'00'

Date (ddmmmyyyy):

3 TABLE OF CONTENTS

4 ABBREVIATIONS

Abbreviation	Description
DET	Detection test
HVLT-R	Hopkins Verbal Learning Test-Revised
IDN	Identification test
ISLT	International Shopping List test
OCL	One Card Learning test
ONB	One Back Memory test
SAP	Statistical Analysis Plan
SAS	Statistical Analysis System
SD	Standard Deviation
SPP	Statistical Programming Plan

5 INTRODUCTION

This Statistical Analysis Plan (SAP) contains technical and detailed elaboration of the principal features of the analysis described in the protocol related to computerized cognitive battery and HVLT-R and includes detailed procedures for executing the statistical analysis of the data.

The SAP will be finalized and signed prior to database lock. If needed, revisions to the approved SAP may be made prior to database lock. Revisions will be version controlled.

This statistical analysis is coordinated by the responsible Cogstate statistician. Any changes from the analyses planned in the SAP will be justified in the Cogstate statistical report.

6 VISIT SCHEDULE

Scheduled Visits

Table 1: Scheduled visits related to Cogstate battery and HVLT-R

Period	Visit Number	Study Day	Study Week by phase	Computerized test battery & HVLT-R Assessment Included
Screening/prospective Observation Phase	**1.2**	**-**	**2**	**Computerized test battery Practice Session**
Double-blind Induction Phase	**2.1**	**1**	**1**	**Yes**
	2.10 or 2.9	**28**	**4**	**Yes**
	EW [b]	**EW**	**-**	**Yes**
Follow-up Phase	**3.2**	**-**	**2 after last intranasal dose**	**Yes**
Follow-up Phase	**3.6 to 3.x**	**Every 84 days**	**Every 12 Weeks**	**Yes for ESKETINTRD3004**

[b] If a subject withdraws before the end of the double-blind induction phase (i.e., before completing Visit 2.10/Day 28) for reasons other than withdrawal of consent, an early withdrawal visit should be conducted within 1 week of the date of discontinuation, followed by the follow-up phase. If the early withdrawal visit is conducted on the same day as a scheduled visit, duplicate assessments are not required.

7 STUDY OBJECTIVES RELEVANT TO COGSTATE ANALYSIS

The objective is to assess the effect of intranasal esketamine on cognition.

8 STUDY DESIGN

The study designs for each study are described in their respective protocols.

9 SAMPLE SIZES

The sample sizes for each study are as follows:

- *116 subjects per treatment for ESKETINTRD3001.*
- *98 subjects per treatment for ESKETINTRD3002.*
- *A total of 211subjects for ESKETINTRD3003.*
- *There is no formal sample size calculation for ESKETINTRD3004 study (Note: the total number of subjects will be based on subjects from this study and subjects from other intranasal esketamine Phase 3 studies).*
- *74 subjects per treatment for ESKETINTRD3005.*

10 ANALYSIS SETS

Safety Analysis Set: All randomized subjects who receive at least 1 dose of intranasal study medication or 1 dose of oral antidepressant medication in the double-blind induction phase. This analysis set will be used for ESKETINTRD3001/2/5.

For ESKETINTRD3003, the safety analysis set for each phase is defined as all subjects who receive at least 1 dose of intranasal study drug or 1 dose of oral antidepressant during that phase.

***Full Analysis Set for open-label induction phase:** will be defined as all subjects who receive at least one dose of intranasal esketamine or 1 dose of oral antidepressant during this phase.*
***Full Analysis Set for optimization/maintenance phase:** will be defined as all subjects who receive at least one dose of intranasal esketamine or 1 dose of oral antidepressant during this phase.*

These analysis sets will be used for ESKETINTRD3004

11 ANALYSIS VARIABLES

11.1 Computerized cognitive battery and Hopkins Verbal Learning Test-Revised (HVLT-R) Endpoints

The endpoints for computerized cognitive battery (DET, IDN, ONB, OCL, and GML) and HVLT-R are the scores and change from baseline (Day 1 prior to randomization) scores at each scheduled post baseline time point.

11.2 The computerized tests from the Cogstate Battery.

11.2.1 Summary of the Cogstate Battery

Detection (DET; Psychomotor Function)

The Detection test is a measure of psychomotor function and uses a well-validated simple reaction time paradigm with playing card stimuli. In this test, the playing cards all depict the same joker. The subject is asked to press the **Yes** key as soon as the card in the center of the screen turns face up. The software

measures the speed and accuracy of each response. The test terminates when a subject has correctly responded to 35 trials. The time to respond (to a maximum) is recorded for each trial.

Duration of test 3 minutes

Identification (IDN; Attention)

The Identification test is a measure of visual attention and uses a well-validated choice reaction time paradigm with playing card stimuli. In this test, the playing cards are all either red or black jokers. The subject is asked whether the card displayed in the center of the screen is red. The subject responds by pressing the **Yes** key when the joker card is red and **No** when it is black. The software measures the speed and accuracy of each response. The time to respond (to a maximum) is recorded for each trial. Wrong responses are counted but do not have an effect on the number of correct responses required for the test to terminate.

Duration of test: 2 minutes

One Card Learning (OCL; Visual Learning)

The One Card Learning test is a measure of visual learning and uses a well-validated pattern separation paradigm with playing card stimuli. In this test, the playing cards are identical to those found in a standard deck of 52 playing cards (without the joker cards). The subject is asked whether the card displayed in the center of the screen was seen previously in this test. The subject responds by pressing the **Yes** or **No** key. The software measures the speed and accuracy of each response. Because no card has been presented yet, the first response is always "No". Eighty trials are displayed during this test.

Duration of test 3 minutes

One Back (ONB; Working Memory)

The One Back test is a measure of working memory and uses a well-validated n-back paradigm with playing card stimuli. In this test, the playing cards are identical to those found in a standard deck of 52 playing cards (without the joker cards). The subject is asked whether the card displayed in the center of the screen is the same as the card presented immediately previously. The subject responds by pressing the **Yes** or **No** key. Because no card has been presented yet on the first trial, a correct first response is always **No**. The software measures the speed and accuracy of each response. The time to respond (to a maximum) is recorded for each trial. Wrong responses are counted but do not have an effect on the number of correct responses required for the test to terminate.

Duration of test: 4 minutes

The Groton Maze Learning test (GML; Executive Function)

The Groton Maze Learning test is a measure of problem solving and reasoning and uses a well-validated maze learning paradigm. In this test, the subject is shown a [10 x 10] grid of boxes on a computer screen. A [28]-step pathway is hidden among these [100] possible locations. Each box represents move locations, and the grid refers to the box array (i.e., [10 × 10]). Subjects are required to find the hidden pathway guided by [four] search rules. These rules are: do not move diagonally, do not move more than one box (i.e., do not jump), do not move back on the pathway, and (return to the last correct location after an error). At each step only the most recently selected box is shown. Feedback is given with visual and auditory cues (green check marks and red crosses) to indicate whether the selected box is correct or incorrect. The head of path, or the last correct location, flashes with a green check when two errors are made in succession to indicate to the subject that they must return to this location. [A delayed recall condition is available for this test and requires the subject to find the hidden pathway after a 10-30 minute delay]. There are [20] well-matched alternate pathways available. The software records each move as an error or as a correct move.

Duration of test: 7 minutes

11.2.2 The outcome measures for the Cogstate battery

Although each of these cognitive tests yields multiple outcome measures, research by Cogstate has identified a set of measures that are optimal for the detection of cognitive change in clinical trials at both the group and individual level (Faletti et al., 2006; Maruff et al., 2009; Bland & Altman, 1996).

For each cognitive test, a single primary outcome measure was selected prior to data analysis from each test in the battery. Each primary outcome measure was selected because it has been shown to be optimal for the detection of change:

a) it is drawn from a data distribution that contains only a small probability of floor or ceiling effects and no restriction in the range of possible performance values (Faletti et al., 2006; Bland & Altman, 1996).

b) it is drawn from a distribution that is distributed normally or which can be corrected to normal through the use of appropriate mathematical transformation (e.g., logarithmic base 10, or arcsine) (Faletti et al., 2006; Bland & Altman, 1996).

Table 1 below summarizes the outcome measures for the Cogstate battery, with the tests from which they were derived, the operational definition, and the variable code.

Table 2: Cogstate tests Administered in ESKETINTRD studies, the Cognitive Domains they Assess, and their Primary Outcome Measures

Cogstate test	Cognitive Domain	Primary Outcome Measure	Interpretation of Primary Outcome Score
Detection test(DET)	Attention (simple reaction time)	Speed of performance (mean of the log10 transformed reaction times for correct responses)	Lower score = better performance
Groton Maze Learning test(GML)	Executive Function	Number of errors across all learning trials	Lower score = better performance
Identification test (IDN)	Attention (choice reaction time)	Speed of performance (mean of the log10 transformed reaction times for correct responses)	Lower score = better performance
One Card Learning test (OCL)	Visual Learning	Accuracy of performance (arcsine square root proportion correct)	Higher score = better performance
One Back test (ONB)	Working Memory	Speed of performance (mean of the log10 transformed reaction times for correct responses)	Lower score = better performance

11.2.3 Data Quality Assurance

Data from the Cogstate battery will be collected on computers at all sites and uploaded to the Cogstate database for processing. Cogstate data management staff will query any data discrepancies. Queries will be confirmed and resolved with the sponsor.

11.2.4 Test Completion Criteria

For each of the Cogstate tests, subjects must provide sufficient responses to allow computation of reliable performance measures. For the majority of Cogstate tests, the term "sufficient" has been defined as a test Completion criterion. The number of trials required for test Completion is unique to each test. They do not vary for different patient groups or study samples.

The completion criteria set forth a priori for each test were as follows:

- DET: The number of responses provided by the subject is ≥ 75% of the desired number of trials (responses ≥ 27).
- IDN: The number of responses provided by the subject is ≥ 75% of the desired number of trials (responses ≥ 23).
- ONB: The number of responses provided by the subject is ≥ 75% of the desired number of trials (responses ≥ 24).
- OCL: 75% of the desired numbers of trials were displayed to the subject (trials ≥ 60).

- GML: The subject provided 28 correct moves in each of the 5 learning trials.

11.3 The HVLT-R

The HVLT-R, a measure of verbal learning and memory, is a 12-item word list recall test. Administration includes 3 learning trials, a 24-word recognition list (including 12 target and 12 foil words), and a delayed recall (20-minute) trial. Administration is computer-assisted; instructions and word lists appear on-screen. The test administrator records each word correctly recalled, and scores for learning, short-term, and delayed recall are generated via the test software. The HVLT-R is a well-validated and widely used measure of verbal episodic memory. The tests will be administered in the following order: HVLT-R, computerized cognitive test battery, and HVLT-R Delayed.

12 STATISTICAL METHODOLOGY

12.1 Analysis Overview

All statistical analyses and summary information will be generated according to this Statistical Analysis Plan (SAP).

For continuous variables, descriptive statistics will include number of subjects (n), mean, standard deviation, median, minimum and maximum. Frequencies and percentages will be displayed for categorical data.

Listings will be produced and displayed.

12.1.1 Analysis of Cogstate Endpoints and HVLT-R

Descriptive statistics (n, mean, SD, min, median, max) of scores and change from baseline scores will be summarized and plotted at each scheduled time point.

For HVLT-R, Scores and change from baseline scores will include the following:
- Total Recall (sum of total correct responses for Trials 1, 2, & 3),
- Delayed Recall (Trial 4)
- Total number of true-positive errors and
- Recognition Discrimination Index (Total no. of true-positives)-(Total no. of false-positives).

12.1.2 Baseline Definition

Baseline is defined as Induction phase, Day 1.

12.1.3 Handling of Missing Data and data transformation

Missing Data

No imputations will be performed in the event of missing data due to dropouts or omitted visits. All incomplete subject profiles, consisting of time points that passed Test completion criteria, will be retained in the analysis. In view of issues of reliability, based on the recommendation of Cogstate scientists, all analyses will be conducted with values which failed completion criteria (listed in Section 11.2.3) removed.

Data Transformation

Speed outcome measure

Since Cogstate speed outcome measures are skewed to the right, $\log_{10}$ transformation will be used to normalize the data. The transformed data will be used in the analyses which are specified in this SAP:

LMN = mean (log10 transformed reaction times for correct responses) (1)

Proportion of accuracy outcome measure

Cogstate correct responses follow a binomial distribution. Since a binomial distribution is not well approximated by a normal distribution, especially when n is small, arcsine square root transformation (which is a variance stabilizing transformation) will be used on Cogstate data on the proportion of correct responses measure and this transformed data will be used in the analysis which is specified in this SAP:

ACC = arcsine (sqrt [proportion of correct responses]) (2)

12.2 Unblinding Procedure

Once all data discrepancies within the Cogstate database are resolved with the clinical research units, the database will be locked and Cogstate will receive the randomization codes from the sponsor.

Table, listing and figure (TLF) shells will be generated in a separate document (Statistical Programming Plan<SPP>).

13 REFERENCES

1. Falleti M.G., Maruff P., Collie A., Darby D.G., & McStephen M. (2003). Qualitative similarities in cognitive impairment associated with 24h of sustained wakefulness and a blood alcohol concentration of 0.05%. Journal of Sleep Research, 12, 265-274.

2. Falleti M.G., Maruff P., Collie A., & Darby D.G. (2006). Practice effects associated with the repeated session of cognitive function using the CogState battery at 10- minute, one week and one-month test retest intervals. Journal of Clinical and Experimental Neuropsychology, 28, 1095-1012.

3. Maruff P., Thomas E., Cysique L., Brew B., Collie A., Snyder P., Pietrzak RH (2009). Validity of the CogState brief battery: Relationship to standardized tests and sensitivity to cognitive impairment in mild traumatic brain injury, schizophrenia, and AIDS dementia complex. Archives of Clinical Neuropsychology, 24, 165-178.