Janssen Research & Development

Statistical Analysis Plan

A Randomized, Double-blind, Multicenter Active-controlled Study to Evaluate the Efficacy, Pharmacokinetics, Safety and Tolerability of Flexible Doses of Intranasal Esketamine Plus an Oral Antidepressant in Adult Subjects with Treatment-resistant Depression

Protocol ESKETINTRD3006; Phase 3
AMENDMENT 2

JNJ-54135419 (esketamine)

Status:	Approved
Date:	5 February 2021
Prepared by:	Janssen Research & Development, LLC
Document No.:	EDMS-ERI-163890143

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

AMENDMENT HISTORY

SAP Version	Issue Date
Original SAP	May 8, 2019
Amendment 1	May 25, 2020
Amendment 2	February 2, 2021

Amendments below are listed beginning with the most recent amendment.

Amendment 2 (02 Feb 2020)

The overall reason for the amendment: to change the first key secondary endpoint according to protocol amendment 2.

Applicable Section(s)	Description of Change (s)
Section 1.1 Trial Objectives	Change the first key secondary objective to 'Change from baseline in the MADRS total score at 24 hours post first dose'. The current first key secondary endpoint is changed to an "Other Secondary Endpoint".
Section 2.1 Visit Windows	For Pulse oximetry, 'End point(DB)' is removed since Pulse ox is only done on dosing days.
Section 2.2.5 Pharmacokinetics analysis Set	The definition of PK analysis set is added.
Section 2.3 Definition of subgroup	History of hypertension is removed. Based on clinical experience, it is not an important factor for the efficacy.
Section 4.1 Demographics and Baseline Characteristics	'Duration of current episode' and 'Duration of current episode categorize (<=2 years, >2 years)' are added in psychiatric history at baseline variables.
Section 5 Efficacy	Key secondary endpoint is updated accordingly in table 5 (efficacy variables) and fixed sequence approach.
Section 5.2.2	Estimand is updated based on latest version of ICH E9
Section 5.3 Major Secondary Endpoints	Change the first key secondary objective to 'Change from baseline in the MADRS total score at 24 hours post first dose'. And use the same methods (MMRM) as the primary endpoint.
Section 5.4 Other Efficacy Variables	The current first key secondary endpoint (Onset of clinical response) is changed to an "Other Secondary Endpoint".
Section 5.4.6 GAD-7	Analysis is added according to protocol
Section 6.1 Adverse Events	AE of special interest has been updated to follow global study TRD3008
Section 6.3 Vital Signs, Weight and BMI and Section 6.5.1 Nasal Examination	Minor corrections are made: The information for hypertension status is updated based on current eCRF Phase information is corrected based on protocol

Amendment 1 (25 May 2020)

The overall reason for the amendment:

Applicable Section(s)	Description of Change (s)
Section 1 Introduction	Since the biomarker application was rejected by the Human Genetic Resource Administration Office of China (HGRAO), biomarker samples will be collected for subjects from the US sites only. The biomarker analysis plan will be developed based on the availability of sufficient number of subjects.
Section 1 Introduction:	General description of potential analysis due to impact of COVID-19 is added.
Section 2.1 Visit Windows	The time window for a visit during the follow-up phase has been modified relative to the last intranasal dose of study drug, instead of the start date of the follow-up phase, to better reflect the Time & Event schedule within the protocol.
Section 2.3 Definition of subgroups	Duration of current episode， Number of classes of prior antidepressant medications and Individual oral drug (duloxetine, escitalopram, venlafaxine XR, sertraline) have been added in order to better understand the depression history of the subject in interpreting the data.
Section 2.4 Study Day and Relative Day	Clarified the purpose of using two relative days for the follow-up phase, one relative day to assign the time window relative to the date of last intranasal dose of study drug and the other relative day relative to the start date of follow-up phase to be presented in the listing.
Section 4.1 Demographics and Baseline Characteristics	Number of depressive episodes, Duration of Current Episode and Number of oral antidepressant drug classes have been added to better present baseline characteristics.
Section 5.4 Other efficacy Variables	By country efficacy analyses have been added for the MADRS total score (including responders, remitters), CGI-S, and SDS analysis, in order to improve on data presentation between subjects in China and the US
Section 5.4.2 Improvement of at least 30% in the MADRS Total Score	An improvement at least 30% in the MADRS total score from baseline has been added as an exploratory endpoint in order to better understand the data
Section 5.4.9 Six-item MADRS subscale	The six-item MADRS subscale has been added as exploratory endpoint in order to better understand data.
Section 6.1 Adverse events	By country safety analysis has been added for the adverse event summary tables to better display data.
Section 6.5.3 Clinician Administered Dissociative States Scale (CADSS)	By country analyses have been added for CADSS in order to better display data.

ABBREVIATIONS

AD	antidepressant
AE	adverse event
ALT/SGPT	alanine aminotransferase
AST/SGOT	aspartate aminotransferase
ATC	anatomic and therapeutic Class
AUC	area under the curve
BMI	body mass index
CADSS	Clinician Administered Dissociative States Scale
CGI-S	Clinical Global Impression –Severity
CMH	Cochran-Mantel-Haenszel
Cmax	maximum concentration
CRF	case report form
CSR	Clinical Study Report
C-SSRS	Columbia-Suicide Severity Rating Scale
CV	coefficient of variation
DB	Double-blind
ECG	electrocardiogram
eCRF	electronic case report form
EQ-5D-5L	EuroQol Group; 5 dimension; 5 level
EQ-VAS	EuroQol Group: visual analogue scale
FAS	full analysis set
FDA	Food and Drug Administration
F/U	Follow-up
GAD-7	Generalized Anxiety Disorder 7-item scale
ICH	International Conference on Harmonization
IVRS	interactive voice response system
LLOQ	lower limit of quantification
MADRS	Montgomery-Asberg Depression Rating Scale
MCMC	Markov Chain Monte Carlo
MDD	major depressive disorder
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple imputation
MINI	Mini International Neuropsychiatric Interview
MGH-ATRQ	Massachusetts General Hospital Antidepressant Treatment Response Questionnaire
MMRM	Mixed-effects model using repeated measures
MRD	minimum required dilution
PD	pharmacodynamic
PK	pharmacokinetic(s)
PP	per protocol
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
SNRI	Serotonin and Norepinephrine Reuptake Inhibitors
SSRI	Selective Serotonin Reuptake Inhibitors
TEAE	treatment-emergent adverse event
TEMA	treatment-emergent markedly abnormal
TRD	Treatment Resistant Depression
ULN	Upper limit of normal

1. INTRODUCTION

This statistical analysis plan (SAP) contains definitions of analysis sets, derived variables, and statistical methods for all planned analyses for study JNJ54135419-ESKETINTRD3006. Pharmacokinetic analyses using population PK modeling will be provided in a separate plan. In addition, if the available number of subjects is sufficient for analysis, the analysis plan for biomarker and pharmacogenomic analyses will be provided separately.

In considering the COVID-19 pandemic has potential impact on the study outcomes, additional analysis may be implemented if applicable.

1.1. Trial Objectives

Primary Objective

The primary objective of this study is to evaluate the efficacy of switching adult subjects with TRD from a prior antidepressant treatment (to which they have not responded) to flexibly dosed intranasal esketamine (56 mg or 84 mg) plus a newly initiated oral antidepressant, compared with switching to a newly initiated oral antidepressant (active comparator) plus intranasal placebo, in improving depressive symptoms, as assessed by the change from baseline in the Montgomery-Asberg Depression Rating Scale (MADRS) total score from Day1 (pre-randomization) to the end of the 4-week double-blind treatment phase.

Key Secondary Objectives

The key secondary objectives are to assess the effect of intranasal esketamine plus a newly initiated oral antidepressant compared with a newly initiated oral antidepressant (active comparator) plus intranasal placebo on the following parameters in adult subjects with TRD:

- Change from baseline in the MADRS total score at 24 hours (Day 2) post first dose
- Functioning and associated disability

Other Secondary Objectives

To assess the effect of intranasal esketamine plus a newly initiated oral antidepressant compared with a newly initiated oral antidepressant (active comparator) plus intranasal placebo on the following parameters in adult subjects with TRD:

- Onset of clinical response by Day 2
- Depression response rates
- Depression remission rates
- Overall severity of depressive illness
- Anxiety symptoms
- Health-related quality of life and health status

To investigate the safety and tolerability of intranasal esketamine plus a newly initiated oral antidepressant compared with a newly initiated oral antidepressant (active comparator) plus intranasal placebo in adult subjects with TRD, including the following parameters:

- Treatment-emergent adverse events (TEAEs), including AEs of special interest
- Local nasal tolerability
- Effects on heart rate, blood pressure, respiratory rate, and blood oxygen saturation
- Dissociative symptoms
- Potential effects on suicidal ideation/behavior
- Potential withdrawal and/or rebound symptoms following cessation of intranasal esketamine treatment.

To assess the pharmacokinetics (PK) of intranasal esketamine in Chinese adult subjects with TRD receiving intranasal esketamine plus a newly-initiated oral antidepressant.

Exploratory Objectives

For subjects who receive esketamine during the double-blind treatment phase, the proportion of subjects who relapse during the follow-up phase will be evaluated separately for subjects who are remitters (MADRS ≤12) at the end of the 4 week double-blind treatment phase and for subjects who are responders but not remitters (≥50% reduction in MADRS total score, but MADRS >12) at the end of the 4 week double-blind treatment phase.

- Time to relapse during the follow-up phase will be evaluated in esketamine subjects who remitted
- Time to relapse during the follow-up phase will be evaluated in esketamine subjects who are responders but not remitters

To assess the potential relationship of biomarkers with response/non-response to intranasal esketamine or oral antidepressants in adult subjects with TRD[a].

[a]: Not applicable to China

1.2. Trial Design

This is a randomized, double-blind, active-controlled, multicenter study in male and female adult subjects with TRD to assess the efficacy, safety, tolerability, and pharmacokinetics of flexibly dosed intranasal esketamine (56 mg or 84 mg) plus a newly initiated oral antidepressant compared with a newly initiated oral antidepressant (active comparator) plus intranasal placebo.

The study has 3 phases;

- Screening/Prospective Observational Phase
- Double-blind Treatment Phase
- Follow-up Phase

Screening/Prospective Observational phase (4-week duration + optional 3-week taper period)

This phase will prospectively assess treatment response to the subject's current oral antidepressant treatment regimen.

At the start of the screening/prospective observational phase, the subject must have had documented non-response to at least 1 antidepressant treatment (based on Massachusetts General Hospital–Antidepressant Treatment Response Questionnaire [MGH-ATRQ]) in the current episode of depression, and subject is taking a different oral antidepressant treatment (on the MGH-ATRQ) for at least the previous 2 weeks at or above the minimum therapeutic dose. This antidepressant treatment, as well as any other ongoing medications being taken for depression at screening (including adjunctive/augmentation therapies), will continue from the start of Week 1 through the end of Week 4 of the screening/prospective observational phase. Dose adjustment is permitted per clinical judgment, but the oral antidepressant treatment is to remain at or above the minimum therapeutic dose (per the MGH-ATRQ) through the end of Week 4. Antidepressant treatment adherence will be assessed using the PAQ. Subjects who report missing $\geqslant$4 days of antidepressant medication treatment(s) in the prior 2-week period will be considered as screen failed due to inadequate adherence.

After 4 weeks, subjects who are non-responders to their current oral antidepressant treatment (as assessed by independent, remote raters) may be eligible to proceed to the double-blind treatment phase. Non-response at the end of the screening/prospective observational phase is defined as $\leqslant$25% improvement in the MADRS total score from Week 1 to Week 4 and a MADRS total score of $\geqslant$28 on Week 2 and Week 4.

Eligible subjects who are entering the double-blind treatment phase will discontinue all of their current medication(s) being taken for depression (including adjunctive/augmentation therapies), and any other prohibited psychotropic medications (including adjunctive atypical antipsychotics). 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, eszopiclone and ramelteon) during the screening/prospective observational phase, can continue these medications during the treatment phase. No dose increases beyond the equivalent of 6 mg/day of lorazepam or new benzodiazepine medications are permitted during the screening/prospective observational phase. If clinically indicated, a subject's current antidepressant treatment(s) may be tapered and discontinued over an additional, optional period of up to 3 weeks per the local prescribing information or clinical judgment, or discontinued and switched directly to one of the four new oral antidepressant medications on Day 1 of the double-blind treatment phase, per clinical judgment.

Double-blind Treatment Phase (4-week duration)

Approximately 234 eligible subjects (210 Chinese subjects, 24 non-Chinese subjects) with TRD, will be randomly assigned at a 1:1 ratio (n=117 subjects per treatment arm, with approximately 105 Chinese subjects per arm) to receive double-blind treatment with either intranasal esketamine or intranasal placebo. The intranasal treatment sessions (esketamine 56 mg, 84 mg or placebo) will

occur twice per week for 4 weeks as a flexible dose regimen at the study site. In addition, all subjects will initiate a new open-label oral antidepressant on Day 1, that will be taken daily for the duration of the double-blind treatment phase. 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 non-response to in the current depressive episode, has not been previously intolerant to (lifetime), and is available in the participating country.

Subjects who were taking benzodiazepines (at dosages equal to or less than the equivalent of 6mg/day of lorazepam) and/or permitted non-benzodiazepine sleep medications (eg, zolpidem, zaleplon, eszopiclone and ramelteon) during the screening/prospective observational phase can continue these medications during the double-blind treatment phase. No dose increases beyond the equivalent of 6 mg/day of lorazepam or new benzodiazepine medications are permitted during the double-blind treatment phase, with the exception of the use of permitted benzodiazepine rescue medication (eg, alprazolam, clorazepate, triazolam, lorazepam and temazepam).

Intranasal Treatment Sessions:

All subjects will self-administer the intranasal study medication (esketamine or placebo) at treatment sessions occurring twice a week for 4 weeks at the clinical site. Treatment is assigned by the Interactive Web Response System (IWRS).

On Day 1, subjects randomized to intranasal esketamine will start with a dose of 56 mg. On any of the subsequent dosing days (Day 4, 8, 11, 15, 18, 22 and 25), the investigator can judge, based on efficacy and tolerability, whether to increase the dose of intranasal esketamine to 84 mg or to maintain the dose at 56 mg. In the event that an increase in dose is poorly tolerated, the investigator may decide to reduce the dose to 56 mg.

Oral Antidepressant Treatment:

Starting on Day 1 of the double-blind treatment phase, a new open-label oral antidepressant treatment will be initiated in all subjects. The oral antidepressant selected will be 1 of the 4 oral antidepressant medications allowed (duloxetine, escitalopram, sertraline, or venlafaxine extended release [XR]). The antidepressant medication will be assigned by the investigator based on review of MGH-ATRQ and relevant prior antidepressant medication information, and will be one that the subject has not previously had a non-response to, in the current major depressive episode, has not been previously intolerant to (lifetime), and is available in participating country.

The use of the titration schedule provided in the protocol is mandatory. Doses are not to exceed the maximum dose defined in the titration schedule. If higher doses are not tolerated, a down-titration is permitted based on clinical judgment.

If a subject withdraws from the study before the end of the double-blind treatment phase for reasons other than withdrawal of consent, an Early Withdrawal visit should be conducted within 1week of the date of discontinuation, followed by the follow-up phase.

Follow-up Phase (8 weeks or until relapse)

The follow-up phase will include all subjects who received at least 1 dose of intranasal study medication in the double-blind treatment phase unless the subjects will directly roll over to study 54135419TRD3008 after completing the 4-week treatment phase. No intranasal study medication will be administered during this phase. Subjects participating in the ESKETINTRD3006 study, at US sites only, following completion of the 4-week double-blind treatment phase, and based on the clinical judgement of the investigator may proceed directly to the 54135419TRD3008 open label extension safety study, without entering the follow-up phase of the ESKETINTRD3006 study. Please refer to the 54135419TRD3008 protocol for full details.

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 and medication changes are permitted. The decision to continue the oral antidepressant in this phase will be at the discretion of the investigator.

This follow-up phase will provide additional information required to assess the course of the subject's major depressive episode over an 8-week period. Safety and tolerability, including potential withdrawal symptoms, following discontinuation of intranasal esketamine will also be assessed.

The follow-up phase is completed at the end of the 8-week period or when a subject meets the criteria below for relapse (in remitters and responders), whichever occurs first.

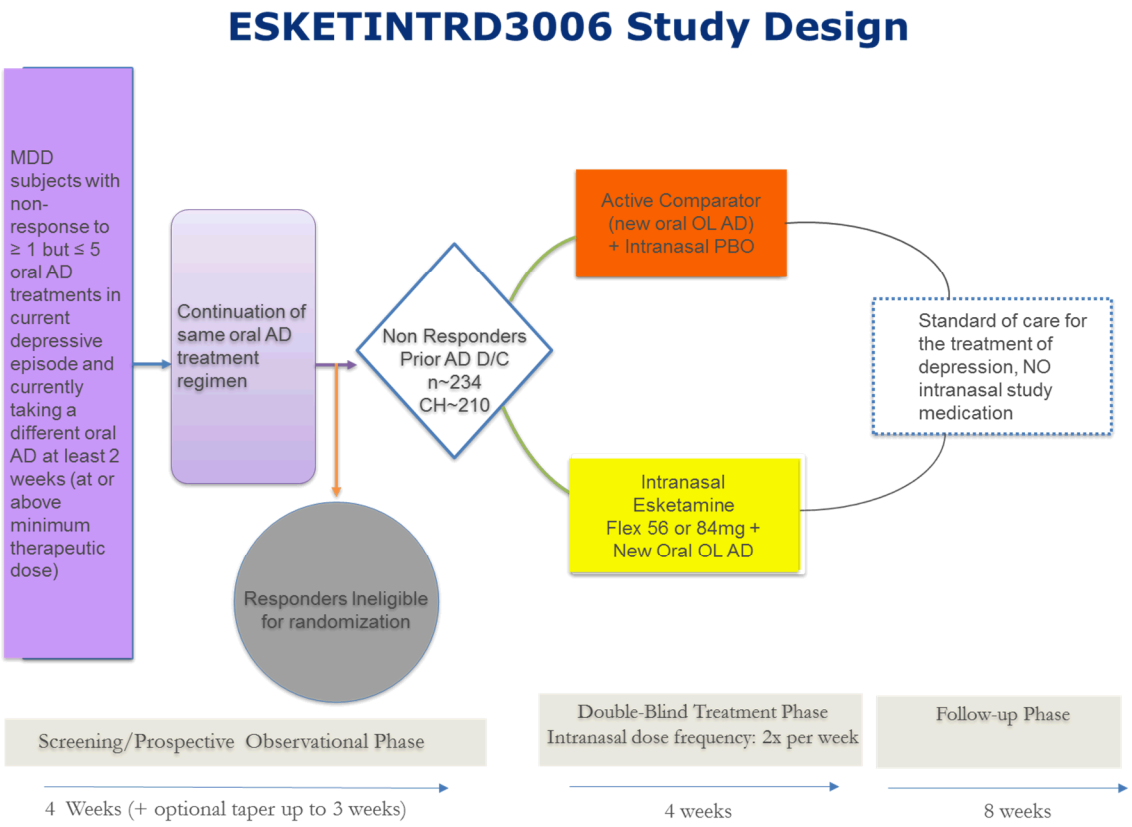
Relapse is defined as any of the following:

- MADRS total score ≥22 for 2 consecutive assessments separated by 5 to 15 days. The date of the second MADRS assessment will be used for the date of relapse.
- Hospitalization for worsening depression or any other clinically relevant event determined per clinical judgment to be suggestive of a relapse of depressive illness such as suicide attempt, completed suicide, or hospitalization for suicide prevention. If hospitalized for any of these events, the start date of hospitalization will be used for the date of relapse. Otherwise the date of the event will be used if the subject is not hospitalized.
- In case more than one relapse criteria are met, the earlier date will be defined as the date of relapse for this subject.

The total duration of a subject's study participation will be up to 19 weeks (including an optional 3-week taper period) for subjects completing the 8-week follow-up phase.

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

Figure 1: Schematic Overview of the Study



- AD = antidepressant; CH = Chinese subjects; D/C = discontinue; MDD = major depressive disorder;
- OL = open-label; PBO = placebo.
- The study will end after 8 weeks follow-up clinical/standard of care for treatment of depression, or following relapse, whichever is earlier.

1.3. Statistical Hypotheses for Trial Objectives

The primary efficacy endpoint is the change in the MADRS total score as measured by the change from baseline (Day 1 prior to randomization) to the end of the 4-week double-blind treatment phase. The hypothesis for this study is that, in adult subjects with TRD, switching from a failed antidepressant treatment to intranasal esketamine plus a newly initiated oral antidepressant is superior to switching to a newly initiated oral antidepressant treatment (active comparator) plus intranasal placebo in improving depressive symptoms based on the primary efficacy endpoint.

1.4. Sample Size Justification

Assuming a treatment difference for the double-blind phase of 6.5 points in MADRS total score between esketamine plus oral antidepressant compared with oral antidepressant plus intranasal placebo at the Week 4 endpoint, a standard deviation of 12, a one-sided significance level of 0.025, and a drop-out rate of 25%, approximately 117 subjects will be randomized to each treatment arm to achieve greater than 90% power, and a sample size of 105 Chinese subjects per treatment arm (total n = 210) plus 12 non-Chinese subjects per treatment arm (total n = 24) will be included. The treatment difference and standard deviation used in this calculation were based on results from Panel A of the ESKETINTRD2003 study and based on clinical judgment.

1.5. Randomization and Blinding

Randomization will be used to minimize bias in the assignment of subjects to treatment groups, to increase the likelihood that known and unknown subject attributes (eg, demographic and baseline characteristics) are evenly balanced across treatment groups, and to enhance the validity of statistical comparisons across treatment groups.

Central randomization will be implemented in this study. Subjects will be randomly assigned to 1 of 2 treatment groups in a 1:1 ratio based on a computer-generated randomization schedule prepared before the study by or under supervision of the sponsor. The randomization will be balanced by using randomly permuted blocks and will be stratified by country, class of oral antidepressant (SNRI or SSRI) and consent to biomarker evaluation (Yes or No) to be initiated in the double-blind treatment phase. The interactive web response system (IWRS) will assign a unique treatment code, which will dictate the treatment assignment and matching study drug kits for the subject. After the investigator selects the oral antidepressant treatment for the double-blind treatment phase, the site will enter this information into IWRS. The requestor must use his or her own user identification and personal identification number when contacting the IWRS, and will then give the relevant subject details to uniquely identify the subject.

The investigator will not be provided with randomization codes. The codes will be maintained within the IWRS, which has the functionality to allow the investigator to break the blind for an individual subject.

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

Under normal circumstances, the blind should not be broken until all subjects have completed the study and the database is finalized. Otherwise, the blind should be broken only if specific emergency treatment/course of action would be dictated by knowing the treatment status of the subject. In such cases, the investigator may in an emergency determine the identity of the treatment by contacting the IWRS. It is recommended that the investigator contact the sponsor or its designee, if possible, to discuss the particular situation, before breaking the blind. Telephone contact with the sponsor or its designee will be available 24 hours per day, 7 days per week. In the event the blind is broken, the sponsor must be informed as soon as possible. The date and time of the unblinding will be documented by the IWRS, and reason for the unblinding must be documented by the electronic case report form (eCRF) and in the source document. The documentation received from the IWRS indicating the code break must be retained with the subject's source documents in a secure manner.

Subjects who have had their treatment assignment unblinded should continue to return for scheduled early withdrawal and follow up visits.

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

To maintain the blinding of intranasal study medication, the esketamine and placebo intranasal devices will be indistinguishable. An intranasal placebo control will be used in the double-blind treatment phase to establish the frequency and magnitude of changes in clinical endpoints that may occur in the absence of intranasal active treatment.

2. GENERAL ANALYSIS DEFINITIONS

2.1. Visit Windows

As subjects do not always adhere to the protocol visit schedule, the following rules are applied to assign actual visits to analysis visits. Listed below are the visit windows and the target days for each visit. The reference day is Study Day 1 (definition in section 2.4). If a subject has 2 or more actual visits in one visit window, the visit closest to the target day will be used as the protocol visit for that visit window. The other additional visit(s) will not be used in the summaries or analyses, but they can be used for determination of clinically important endpoints. If 2 actual visits are equidistant from the target day within a visit window, the later visit is used.

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

Table 1: Visit Windows

Parameter	Analysis Phase	Scheduled Visit Number	Time Interval (label on output)	Time Interval[a] (Day)	Target Time[a] Point (Day)
MADRS	Screening	1.1	Week 1 (SC)		
		1.2	Week 2 (SC)		
		1.3	Week 4 (SC)		
	DB	2.1	Baseline	≤1	1
		2.2	Day 2	2-3	2
		2.4	Day 8	4-11	8
		2.6	Day 15	12-18	15
		2.8	Day 22	19-24	22
		2.10	Day 28	25 to end of DB	28
		DB final visit	End Point (DB)	2 to end of DB	
	Follow-up	3.1	Week 1 (F/U)	1- 10	7
		3.2	Week 2 (F/U)	11-17	14
		3.3	Week 3 (F/U)	18-24	21
		3.4	Week 4 (F/U)	25-31	28
		3.5	Week 5 (F/U)	32-38	35
		3.6	Week 6 (F/U)	39-45	42
		3.7	Week 7 (F/U)	46-52	49
		3.8	Week 8 (F/U)	53 to end of F/U	56
		F/U final visit	End point (F/U)	2 to end of F/U	

Table 1: Visit Windows

Parameter	Analysis Phase	Scheduled Visit Number	Time Interval (label on output)	Time Interval[a] (Day)	Target Time[a] Point (Day)
SDS	Screening	1.1	Week 1 (SC)		
	DB	2.1	Baseline	⩽1	1
		2.6	Day 15	12-21	15
		2.10	Day 28	22 to end of DB	28
		DB final visit	End Point (DB)	2 to end of DB	
	Follow-up	3.2	Week 2 (F/U)	1-24	14
		3.5	Week 5 (F/U)	25-40	35
		3.8	Week 8 (F/U)	41 to end of F/U	56
		F/U final visit	End point (F/U)	2 to end of F/U	
CGI-S	Screening	1.1	Week 1 (SC)		
	DB	2.1	Baseline	⩽1	1
		2.3	Day 4	2-6	4
		2.4	Day 8	7-9	8
		2.5	Day 11	10-13	11
		2.6	Day 15	14-18	15
		2.8	Day 22	19-24	22
		2.10	Day 28	25 to end of DB	28
		DB final visit	End Point (DB)	2 to end of DB	
	Follow-up	3.2	Week 2 (F/U)	1-24	14
		3.5	Week 5 (F/U)	25-40	35
		3.8	Week 8 (F/U)	41 to end of F/U	56
		F/U final visit	End point (F/U)	2 to end of F/U	
GAD-7	Screening	1.1	Week 1 (SC)		
	DB	2.1	Baseline	⩽1	1
		2.10	Day 28	2 to end of DB	28
		DB final visit	End Point (DB)	2 to end of DB	
	Follow-up	3.2	Week 2 (F/U)	124	14
		3.5	Week 5 (F/U)	25-40	35
		3.8	Week 8 (F/U)	41 to end of F/U	56
		F/U final visit	End point (F/U)	2 to end of F/U	
EQ-5D-5L	Screening	1.1	Week 1 (SC)		
	DB	2.1	Baseline	1	1
		2.6	Day 15	2-21	15
		2.10	Day 28	22 to end of DB	28
		DB final visit	End Point (DB)	2 to end of DB	
	Follow-up	3.2	Week 2 (F/U)	124	14
		3.5	Week 5 (F/U)	2540	35
		3.8	Week 8 (F/U)	41 to end of F/U	56
		F/U final visit	End point (F/U)	2 to end of F/U	
Vital Signs (TEMP [predose at each visit], BP[b], HR, RESP [at each visit, predose, 40M, 1H, 1.5H])	DB	2.1	Baseline Day 1: Predose Day 1: 40M Day 1: 1H Day 1: 1H30M	≤1/predose	1
		2.3	Day 4: Predose Day 4: 40M Day 4: 1H Day 4: 1H30M	2-6	4
		2.4	Day 8: Predose Day 8: 40M Day 8: 1H Day 8: 1H30M	7-9	8
		2.5	Day 11: Predose Day 11: 40M Day 11: 1H Day 11: 1H30M	10-13	11

Table 1: Visit Windows

Parameter	Analysis Phase	Scheduled Visit Number	Time Interval (label on output)	Time Interval[a] (Day)	Target Time[a] Point (Day)
		2.6	Day 15: Predose Day 15: 40M Day 15: 1H Day 15: 1H30M	14-16	15
		2.7	Day 18: Predose Day 18: 40M Day 18: 1H Day 18: 1H30M	17-20	18
		2.8	Day 22: Predose Day 22: 40M Day 22: 1H Day 22: 1H30M	21-23	22
		2.9	Day 25: Predose Day 25: 40M Day 25: 1H Day 25: 1H30M	24 to end of DB	25
		DB final visit	End Point (DB)	Day 1: 40M to end of DB	
	Follow-up	3.2	Week 2 (F/U)	1-24	14
		3.5	Week 5 (F/U)	25-45	35
		3.8	Week 8 (F/U)	46 to end of F/U	56
		F/U final visit	End point (F/U)	2 to end of F/U	
Weight and BMI	DB	2.1	Baseline	≤1	1
		2.10	Day 28	2 to end of DB	28
		DB final visit	End Point (DB)	2 to end of DB	
	Follow-up	3.8	Week 8 (F/U)	46 to end of F/U	56
		F/U final visit	End point (F/U)	2 to end of F/U	
ECG	DB	2.1	Baseline	≤1/Predose	1
		2.1	Day 1: 1H	1	1
		2.4	Day 8: 1H	2-16	8
		2.9	Day 25: 1H	17 to end of DB	25
		DB final visit	End Point (DB)	2 to end of DB	
Lab (Hematology, Chemistry)	DB	2.1	Baseline	≤1/predose	1
		2.10	Day 28	2 to end of DB	28
		DB final visit	End Point (DB)	2 to end of DB	
	Follow-up	3.2	Week 2 (F/U)	2 to end of F/U	14
Lab (Urinalysis)	DB	2.1	Baseline	≤1/predose	1
		2.6	Day 15	2-21	15
		2.10	Day 28	22 to end of DB	28
		DB final visit	End Point (DB)	2 to end of DB	
	Follow-up	3.2	Week 2 (F/U)	2 to end of F/U	14
C-SSRS	Screening	1.1	Week 1 (SC)		
		1.2	Week 2 (SC)		
		1.3	Week 4 (SC)		
	DB	2.1	Baseline	≤1	1
		2.3	Day 4	2-6	4
		2.4	Day 8	7-9	8
		2.5	Day 11	10-13	11
		2.6	Day 15	14-16	15
		2.7	Day 18	17-20	18
		2.8	Day 22	21-23	22
		2.9	Day 25	24-26	25
		2.10	Day 28	27 to end of DB	28
		DB final visit	End Point (DB)	1 to end of DB	
	Follow-up	3.1	Week 1 (F/U)	1-10	7
		3.2	Week 2 (F/U)	11-17	14
		3.3	Week 3 (F/U)	18-24	21

Table 1: Visit Windows

Parameter	Analysis Phase	Scheduled Visit Number	Time Interval (label on output)	Time Interval[a] (Day)	Target Time[a] Point (Day)
		3.4	Week 4 (F/U)	25-31	28
		3.5	Week 5 (F/U)	32-38	35
		3.6	Week 6 (F/U)	39-45	42
		3.7	Week 7 (F/U)	4652	49
		3.8	Week 8 (F/U)	53 to end of F/U	56
		F/U final visit	End point (F/U)	2 to end of F/U	
PWC-20	DB	2.9	Day 25	2 to end of DB	25
		DB final visit	End Point (DB)	2 to end of DB	
	Follow-up	3.1	Week 1 (F/U)	1-10	7
		3.2	Week 2 (F/U)	11 to end of F/U	14
		F/U final visit	End Point (F/U)	2 to end of F/U	
Nasal exam	DB	2.1	Baseline	≤1	1
		2.10	Day 28	2 to end of DB	28
		DB final visit	End Point (DB)	2 to end of DB	
CADSS	DB	2.1	Day 1: predose Day 1: 40M Day 1: 1H30M	≤1/predose	1
		2.3	Day 4: predose Day 4: 40M Day 4: 1H30M	2-6	4
		2.4	Day 8: predose Day 8: 40M Day 8: 1H30M	7-11	8
		2.6	Day 15: predose Day 15: 40M Day 15: 1H30M	12-18	15
		2.8	Day 22: predose Day 22: 40M Day 22: 1H30M	19-23	22
		2.9	Day 25: predose Day 25: 40M Day 25: 1H30M	24 to end of DB	25
Pulse oximetry (every 15 minutes from predose to t=1.5 hours postdose)[c]	DB	2.1	Day 1 : Predose Day 1: 15M Day 1：30M Day 1: 45M Day 1: 1H Day 1: 1H15M Day 1: 1H30M	≤1/predose	1
		2.3	Day 4 : Predose Day 4: 15M Day 4：30M Day 4: 45M Day 4: 1H Day 4: 1H15M Day 4: 1H30M	2-6	4
		2.4	Day 8: Predose Day 8: 15M Day 8：30M Day 8: 45M Day 8: 1H Day 8: 1H15M Day 8: 1H30M	7-9	8
		2.5	Day 11: Predose Day 11: 15M Day 11：30M Day 11: 45M Day 11: 1H	10-13	11

Table 1: Visit Windows

Parameter	Analysis Phase	Scheduled Visit Number	Time Interval (label on output)	Time Interval[a] (Day)	Target Time[a] Point (Day)
			Day 11: 1H15M Day 11: 1H30M		
		2.6	Day 15: Predose Day 15: 15M Day 15：30M Day 15: 45M Day 15: 1H Day 15: 1H15M Day 15: 1H30M	14-16	15
		2.7	Day 18: Predose Day 18: 15M Day 18：30M Day 18: 45M Day 18: 1H Day 18: 1H15M Day 18: 1H30M	17-20	18
		2.8	Day 22: Predose Day 22: 15M Day 22：30M Day 22: 45M Day 22: 1H Day 22: 1H15M Day 22: 1H30M	21-23	22
		2.9	Day 25: Predose Day 25: 15M Day 25：30M Day 25: 45M Day 25: 1H Day 25: 1H15M Day 25: 1H30M	24 to end of DB	25

[a] For double-blind (DB) treatment phase, time interval is relative to the date of Study Day 1 of the DB phase. For the follow-up (F/U) phase, since the scheduled visits are related to the last intranasal dose, accordingly time interval is relative to the date of the last intranasal dose.
[b] During the DB phase, at 1.5 hours post dose if the SBP is ≥160 mmHg and/or DBP is ≥100 mmHg, assessments should continue every 30 minutes until the blood pressure is <160 mmHg and/or <100 mmHg or investigator's clinical judgment the subject it is clinical stable and can be discharged from the clinical site.
[c] If oxygen saturation levels are <93% at any time during the 1.5 hour postdose interval, pulse oximetry will be recorded every 5 minutes until levels return to ≥93% or until the subject is referred for appropriate medical care, if clinically indicated.

2.2. Analysis Sets

2.2.1. All Randomized Analysis Set

The all randomized analysis set includes all subjects who were randomized in the study (i.e., subjects who reported a randomization date, or were assigned a randomization number) regardless of whether or not treatment was received. This analysis set will be used for summarizing the overall study completion/withdrawal information.

2.2.2. Efficacy Analysis Set(s)

2.2.2.1. Primary Efficacy Analysis Set

Full analysis set (FAS) is the primary efficacy analysis set, which includes all randomized subjects who received a least 1 dose of intranasal study medication and 1 dose of oral antidepressant medication during the double-blind treatment phase. FAS will be used for all efficacy analysis.

2.2.2.2. Secondary Efficacy Analysis Set

Per-protocol set is the secondary efficacy analysis set, which includes a subset of subjects in the FAS who is in-compliance with the protocol. Compliance is defined as no major protocol deviations. List of major protocol deviations is specified in Section 4.5. Per-protocol set will be used as sensitivity analyses for the primary efficacy analyses.

2.2.3. Safety Analysis Set

The safety analysis set is defined for the double-blind treatment phase. The safety analysis set includes all randomized subjects who received at least 1 dose of intranasal study medication or 1 dose of oral antidepressant medication during the double-blind treatment phase. Analyses of change from baseline will include only subjects who have both baseline and at least 1 postbaseline observation during that phase. Screen failures and randomized subjects who received no double-blind study medication will be excluded from the safety analysis set. Subjects who received an incorrect treatment will be analyzed under the planned treatment. This analysis set will be used for safety analyses.

2.2.4. Follow-up Analysis Set

The Follow-up analysis set is defined as all subjects who enter the follow-up phase. This analysis set will be used for both efficacy and safety analyses.

2.2.5. Pharmacokinetics Analysis Set

The pharmacokinetic population will consist of all subjects who received at least 1 dose of intranasal study medication and have at least 1 post treatment sample collected during treatment.

2.2.6. Biomarker and Pharmacogenomics Analysis Set

The definition of analysis set and analyses will be provided in a separate document if the available number of subjects is sufficient.

2.3. Definition of Subgroups

Definition of subgroups are presented in Table 2.

Table 2: Definition of Subgroups

Subgroup	Definition
Country	China US
Age Group	**Adults** • 18-44 years • 45-64 years
Sex	• Female • Male
Number of Previous Treatment Failures in Current Episode	• Based on ATRQ
Functional Impairment based on Baseline SDS Total Score	• No impaired (0-3) • Mile (4-11) • Moderate (12-19) • Marked (20-26) • Extreme (27-30)
Baseline MDRS total Score	• ≤ median • > median
Class of antidepressant study medication	• SSRI • SNRI
BMI	• ≤ median (kg/M2) • > median (kg/M2)
Employee Status	• Employment • Unemployment
Duration of current episode	• ≤ 2 years • > 2 years
Number of classes of prior failed antidepressant medications	• < 2 • ≥ 2
Individual antidepressant study medication	• Duloxetine • Escitalopram • Sertraline • Venlafaxine extended release (XR)

2.4. Study Day and Relative Day

Study Day

Study Day 1 or Day 1 refers to the earlier of date of the first dose of intranasal study medication, or the oral antidepressant study drug (study Day 1 is not available for screened subjects who did not take any intranasal study drug or oral antidepressant study drug). All efficacy and safety assessments at all visits will be assigned a day relative to the date of Study Day 1.

Study day for a visit is defined as:

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

There is no 'Day 0'.

Relative Day

The start date of the double-blind treatment phase (referred to as ‘DB start date’) is the same as date of Study Day 1. The start date/time of the double-blind treatment phase (referred to as, ‘DB start date/time’) is the DB start date and the time of the first dose of intranasal study medication.

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 date of the follow-up phase (referred to as 'F/U start date') is the day after the DB end date.

Relative day to the analysis phase start date for a visit is defined as:

- Visit date - (analysis phase start date) +1, if visit date is ≥ analysis phase start date
- Visit date - analysis phase start date, if visit date is < analysis phase start date

Time window for a follow-up phase is assigned based on the relative day to the date of last intranasal dose, which the relative day is defined as:

- Visit date - (the last intranasal dose date) +1, if visit date is ≥ the last intranasal dose date
- Visit date - the last intranasal dose date, if visit date is < the last intranasal dose date

Note that the relative day to F/U start date will be presented in the listing but not be used to derive the time window of follow-up phase.

The double-blind treatment phase end date (referred to as 'DB end date') is the date of completion/withdrawal from the double-blind treatment 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.

2.5. Baseline and Endpoint

Baseline is defined as the last observation before receiving the first dose of any study medication in the double-blind treatment phase. Baseline is defined for each parameter/assessment.

The double-blind endpoint is defined as the last available postbaseline result during the double-blind treatment phase. Unscheduled visit results are included in this definition and will be considered as the endpoint value if the unscheduled visit result is the last postbaseline result available within the analysis period.

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

Treatment-emergent adverse events (TEAEs) for the double-blind treatment phase are those events with an onset date/time on or after the start of study medication and occurred on or before the end of the double-blind phase. Adverse events (AEs) for the follow-up phase are those events with an onset date on or after the start of the follow-up phase and occurred on or before the end of the follow-up phase. A conservative approach will be used to handle the missing dates for adverse events.

Onset Date

Partial AE onset dates will be imputed as follows:

- If the onset date of an adverse event is missing day only, it will be set to:
  - First day of the month that the AE occurred, if month/year of the onset of AE is different than the month/year of the DB start date
  - The day of DB start date, if the month/year of the onset of AE is the same as month/year of the DB start date and month/year of the AE resolution date is different
  - The day of DB start date or day of AE resolution date, whichever is earliest, if month/year of the onset of AE and month/year of the DB start date and month/year of the AE resolution date are same
- If the onset date of an adverse event is missing both day and month, it will be set to the earliest of:
  - January 1 of the year of onset, as long as this date is on or after the DB start date
  - One day after the DB start date, if this date is the same year that the AE occurred
  - Last day of the year if the year of the AE onset is prior to the year of the DB start date,
  - The AE resolution date.
- Completely missing onset date of an adverse event will be set to the DB start date.

Resolution Date

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

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

Time

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

- A missing time of onset of an adverse event will be set as follows:
  - 00:00 as long as the onset date is after the DB start date
  - 00:00 if the date is the same as DB start date, but the intranasal study medication in the double-blind treatment phase was started after the oral antidepressant medication in this phase

  - The time of intranasal medication start in the double-blind treatment phase if the date is the same as DB start date, and the intranasal medication was started on or before the oral antidepressant medication in this phase.

- The missing time of resolution of an adverse event will be set to 23:59.

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

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:

- If the month of the concomitant medication start date is equal to the month of the start of the Double-Blind phase, then the estimated start date is the DB start date;
- If the month of the concomitant medication start date is greater than the month of the start of the Double-Blind 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 DB 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. INTERIM ANALYSIS AND DATA MONITORING COMMITTEE REVIEW

No interim analysis is planned in this study.

4. SUBJECT INFORMATION

4.1. Demographics and Baseline Characteristics

Demographic and baseline characteristics (Table 3) and psychiatric history at baseline (Table 4) will be summarized by treatment group and overall for the Safety and Full analysis sets. In addition, a separate summary will be presented by treatment group and country (China, non-China) for Safety and Full analysis sets.

Table 3: Demographic Variables

Continuous Variables:	**Summary Type**
Age (years)	Descriptive statistics (N, mean, standard deviation [SD], median and range [minimum and maximum]).
Baseline Weight (kg)	
Baseline Height (cm)	
Baseline Body Mass Index (BMI) (kg/m^2)	
Categorical Variables	
Age (18-44 years and 45-64 years)	Frequency distribution with the number and percentage of subjects in each category.
Sex (male, female, undifferentiated)	
Race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or other Pacific Islander, White, Other, Multiple)	
Ethnicity (Hispanic or Latino, not Hispanic or Latino)	
Baseline BMI ([underweight <18.5 kg/m^2, normal 18.5-<25 kg/m^2, overweight 25-<30 kg/m^2, obese ≥30 kg/m^2])	
Employment Status	
Hypertension Status	
Class of antidepressant (SSRI/SNRI)	
Oral antidepressant	
Country	

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

Table 4: Psychiatric History at Baseline Variables

Continuous Variables:	**Summary Type**
Baseline MADRS total score	Descriptive statistics (N, mean, standard deviation [SD], median and range [minimum and maximum]).
Baseline CGI-S score	
Age (years) when diagnosed with MDD	
Duration of current episode	
Categorical Variables	
Baseline CGI-S score	Frequency distribution with the number and percentage of subjects in each category.
Screening C-SSRS category (no event, suicidal ideation, suicidal behavior)	
Duration of current episode categorize (<=2 years, >2 years)	
Antidepressant treatment history (number of medications with non-response taken for at least 6 weeks during the current episode as obtained in the MGH-ATRQ)	
Number of depressive episodes	
Number of classes of prior failed antidepressant medications	
Family history of - Depression - Anxiety Disorder - Bipolar Disorder - Schizophrenia - Alcohol Abuse - Substance Abuse	

4.2. Disposition Information

The distribution of the number of subjects who are randomized, receive double-blind treatment and complete the double-blind treatment phase will be presented by treatment group for overall population. A separate summary will be presented by treatment group and country (China, US).

In addition, the distribution of reasons for discontinuation will be presented. These summaries will be provided for the All Randomized and Safety analysis sets. A subject will be considered to have completed the double-blind treatment phase if the subject has completed assessment through Day 28. The distribution of the number of subjects who enter the follow-up phase and complete all follow-up visits will be presented by treatment group for China population. The reasons for

discontinuation will be presented. In addition, the number of subjects who enroll in study ESKETINTRD3008 will be provided.

The number of subjects in each analysis set will be presented by treatment group for overall population. The distribution of screened subjects will be presented by country and site ID.

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

A listing of subjects will be provided for the following categories:

- Subjects who discontinued the double-blind treatment phase
- Subjects who discontinued the study
- Subjects who were unblinded during the double-blind treatment phase
- Subjects who were randomized yet did not receive study medication.

4.3. Treatment Compliance

Doses of oral AD will be summarized using descriptive statistics of the mean dose (days on drug), mode dose (days on drug) and the final dose, by each type of oral AD for both the double-blind treatment and follow-up phases. In addition, percent compliance of the oral AD will be calculated as, days actually dosed/days expected to be dosed*100, and summarized.

4.4. Extent of Exposure

Extent of exposure in terms of total duration of exposure and number of dosing sessions of intranasal study medication and oral AD will be summarized for the Full analysis set and the Safety analysis set.

The total duration of exposure for the intranasal study drug and for each type of oral antidepressant (AD) during the double-blind treatment phase is defined as (date of last dose of each type of study medication – date of first dose of each type of study medication) +1.

Descriptive statistics (N, mean, SD, median, minimum and maximum) of total duration of exposure of intranasal study drug will be presented. The total duration of intranasal study drug exposure in the double-blind treatment phase will be presented using the following categories: ≤7 days, 8-14 days, 15-21 days, 22-25 days, >25 days. A frequency distribution of the total number of dosing sessions of intranasal study medication during the double-blind treatment phase will be presented. A frequency distribution of dose level will be presented for each dosing session.

The total duration of exposure of oral AD will be summarized similarly to the intranasal study drug, however the following categories will be used for the double-blind treatment phase: ≤7 days, 8-14 days, 15-21 days, 22-28 days, >28 days and for the follow-up phase: ≤7 days, 8-14 days, 15-21 days, 22-28 days, 29-35 days, 36-42 days, 43-49 days, 50-56 days, >56 days. Each type of oral AD will be summarized separately.

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

Descriptive statistics (N, mean, SD, median, minimum and maximum) of modal dose, mean and final dose of intranasal study drug will be presented. Doses of oral AD will be summarized using descriptive statistics of the mean dose (days on drug), mode dose (days on drug) and the final dose, by each type of oral AD for both the double-blind treatment and follow-up phases.

A summary of descriptive statistics that will be done for the double-blind treatment phase are listed:

- Total duration of exposure of intranasal study drug for double-blind treatment phase
- Total duration of exposure of oral AD for both the double-blind treatment phase and follow-up phase
- Modal dose, mean and final dose of intranasal study drug
- Modal dose, mean and final dose of oral AD

Summary of other analyses include:

- Frequency distribution of the total number of dosing sessions of intranasal study medication during the double-blind treatment phase
- Frequency distribution of dose level will be presented for each dosing session
- Doses of oral AD

4.5. Protocol Deviations

In general, the following list of major protocol deviations may have the potential to impact subjects' rights, safety or well-being, or the integrity and/or result of the clinical study. Subjects with major protocol deviations will be identified prior to database lock and the subjects with major protocol deviations will be summarized by category. More categories may be included depending on the nature of the protocol deviation.

- Developed withdrawal criteria but not withdrawn
- Entered but did not satisfy criteria
- Received a disallowed concomitant treatment
- Received wrong treatment or incorrect dose
- Other

Minor protocol deviations related to COVID-19 will be included in the listing.

4.6. Prior and Concomitant Medications

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

Antidepressant medications taken prior to the baseline visit will be summarized by treatment group for the Safety analysis set.

Summaries of concomitant medications will be presented by treatment group, and study phase for the Safety analysis set and for the Follow-Up analysis set. The proportion of subjects who receive each concomitant medication will be summarized as well as the proportion of subjects who receive at least one concomitant medication.

5. EFFICACY

The efficacy data will be summarized by treatment group in each study phase. Primary and key secondary efficacy analyses will be presented for overall and China populations, separately. Analysis population for other endpoints will be overall population unless otherwise specified.

The efficacy variables for this study are listed in Table 5.

Table 5: Efficacy Variables

Efficacy Variable		**Endpoint**
MADRS	• Change in MADRS from Baseline to Day 28 (DB)	Primary
	• Change in MADRS from Baseline to 24 hours (Day 2)	Key Secondary
	• Onset of Clinical Response (achieving at least 50% improvement with onset by Day 2 that is maintained to Day 28 with one excursion)	Secondary
	• Proportion of responders (≥50% reduction from baseline in MADRS total score) at the end of the 4 week double-blind treatment phase.	Secondary
	• Proportion of subjects in remission (MADRS ≤12) at the end of the 4 week double-blind treatment phase	Secondary
	• Proportion of subjects who remain in remission (MADRS ≤12) through to the end of the follow-up phase	Secondary
	• Proportion of subjects who remain as responders (≥50% reduction from baseline in MADRS total score) through to the end of the follow-up phase	Secondary
	• Response rates over time (at least 50% improvement from baseline)	Secondary
	• Remission rates over time (MADRS≤12)	Secondary
	• Time to sustained response (first occurrence of response that is maintained through Day 28 with one excursion)	Secondary
	• Time to sustained remission (first occurrence of remission that is maintained through Day 28 with one excursion)	Secondary
	• Proportion of subjects ≥30% reduction from baseline in MADRS total score at the end of the 4-week double-blind treatment phase	Exploratory
	• Change in MADRS6 subscale from Baseline to Day 28 (DB)	Exploratory

Table 5: Efficacy Variables

Efficacy Variable		Endpoint
SDS	• Change in SDS from Baseline to Day 28	Key Secondary
	• Change in SDS during the F/U phase	Secondary
	• Response based on SDS over time	Secondary
	• Remission based on SDS over time	Secondary
CGI-S	• Change in CGI-S from Baseline to Day 28	Secondary
GAD-7	• Change in GAD-7 from Baseline to Day 28 or End Point (DB)	Secondary
EQ-5D-5L,	• Change from Baseline to Day 28 or End Point (DB)	Secondary
Relapse	• Time to relapse during the F/U phase for subjects who are remitters at the end of DB phase	Exploratory
	• Time to relapse during the F/U phase for subjects who are responders but not remitters	Exploratory
	• Proportion of subjects who relapse during the F/U phase for subjects who are remitters	Exploratory
	• Proportion of subjects who relapse during the F/U phase for subjects who are responders but not remitters	Exploratory

5.1. Analysis Specifications

5.1.1. Level of Significance

Statistical analysis tests will be conducted at a 2-sided 0.05 level of significance unless specified otherwise.

A serial gatekeeping (fixed sequence) approach will be applied to adjust for multiplicity and to strongly control type I error across the primary and the 2 key secondary efficacy endpoints (onset of clinical response, change in SDS total score. The 2 key secondary endpoints will be analyzed sequentially and will be considered statistically significant at the 2-sided 0.05 level only if the endpoint is individually significant at the 2-sided 0.05 level and previous endpoints in the hierarchy were significant at the 2-sided 0.05 level, including the primary endpoint. If the primary endpoint is statistically significant, the selected secondary endpoints will be assessed in the following order:

- Change in MADRS total score at 24 hours post first dose
- Change in SDS total score from Baseline to Day 28

5.1.2. Data Handling Rules

The last post baseline observation during the double-blind treatment phase will be carried forward as the "End Point" for that phase. Imputation of the MADRS total score 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 left blank.

5.2. Primary Efficacy Endpoint

5.2.1. Definition

The primary efficacy endpoint is the change in MADRS total score from Day 1 to Day 28. The MADRS is a clinician-rated scale designed to measure depression severity and to detect changes due to antidepressant treatment[6]. 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.

5.2.2. Estimand

The primary estimand, the main clinical quantity of interest to be estimated in the study, is defined by the following 5 components:

Treatment: Active comparator (open-label new oral antidepressant) plus intranasal placebo and Intranasal esketamine (56 mg or 84 mg) plus open-label new oral antidepressant

Population: subjects with treatment-resistant depression as defined through appropriate inclusion/exclusion criteria to reflect the targeted patient population

Variable: change from baseline to Day 28 in the MADRS total score

Intercurrent events and corresponding strategies:

- Treatment discontinuation (including COVID-19 related discontinuation) – Hypothetical Strategy: the effect of the initially randomized treatment together with the oral antidepressant medication that would have been observed had all subjects remained on their treatment throughout the double-blind treatment phase
- Major protocol violations and COVID-19 related minor protocol violations – Treatment Policy Strategy: use all measures, regardless of whether or not major protocol violations or COVID-19 related minor protocol violations had occurred.

Population-level summary: the difference in variable least square means

The primary analysis will be based on the full analysis set, as described in Section 2.2.2.1, and the MADRS total scores collected during the DB treatment phase.

5.2.3. Analysis Methods

The primary efficacy variable, change from baseline in MADRS total score at Day 28 in the double-blind treatment phase, will be analyzed using a Mixed-Effect Model for Repeated Measures (MMRM) based on observed case data. The models will include baseline MADRS total score as a covariate, and treatment, country, class of antidepressant (SNRI or SSRI), day (see Table 1), and day-by-treatment interaction as fixed effects. The within-subject covariance between visits will be estimated via an unstructured variance-covariance matrix. In case of convergence problems, alternative variance-covariance structures will be tried in the following order, with the first structure that converges being used in the analysis: heterogeneous Toeplitz, standard Toeplitz, and AR(1) with separate subject random effect. Comparison of the esketamine plus oral antidepressant arm versus oral antidepressant plus intranasal placebo will be performed using the

appropriate contrast. Lease squares mean changes in MADRS total score over time and cumulative percent response for change from baseline to 28 days in MADRS total score will be graphically presented.

Model Diagnostics

The normality and equal variance assumptions will be assessed graphically for the MADRS total score at end point. Residuals from the primary models will be plotted against the predicted values and a QQ plot of the residuals versus the expected quantiles of the standard normal distribution will be presented. If either the equal variance or the normality assumption appears to be grossly violated, other methods like an appropriate transformation of the primary endpoint might be considered.

Missing Data Sensitivity Analysis

The following table (Table 6) shows the assumptions of each considered analysis for the primary efficacy endpoint, all applied to the same full analysis set defined in Section 2.2.2.1:

Table 6: Analysis Types and Assumptions

Analysis Type	Analysis Method	Assumption
Primary Analysis	MMRM	Missing at Random – MAR
Sensitivity Analysis	Delta worsening adjustment applied to standard multiple imputation regression	Efficacy scores worsen after study discontinuation

To evaluate the robustness of the MMRM analysis to increasing deviations from the MAR assumption, a delta adjustment multiple imputation method will be used for sensitivity analysis. This type of method is regarded to be an informative sensitivity analysis in clinical trials (2010 National Research Council report on missing data[7] and T Permutt[8]).

This method will employ the following 3 steps:

Step 1 –Multiple Imputation (MI)

Note: Most missing data will be a result of subjects dropping out of the study and having observations at every scheduled visit up to the point they dropped but no observations thereafter. This is a monotone missing data pattern. Some subjects may have ‘intermediate missing’ data between non-missing observations. This is a non-monotone missing data pattern.

If there are subjects with a non-monotone missing data pattern, datasets with only monotone missing data patterns will be created first by imputing the intermediate missing values using the MCMC (Markov Chain Monte Carlo) method. This will be done using SAS PROC MI and the MCMC statement with the following specifications:

PROC MI DATA=INPUT NIMPUTE=500 SEED=234

OUT=IN_MCMC; VAR …; (treatment group, country, class of antidepressant (SNRI or SSRI), and the preceding non-missing values in the order of clinical visits: baseline, Day 2, Day 8, Day 15, Day 22 and Day 28)

MCMC CHAIN=SINGLE NBITER=200 NITER=100 IMPUTE=MONOTONE;

RUN;

Note: Graphical diagnostic tools available in the MCMC statement (TIMEPLOT(WLF) and ACFPLOT(WLF)) will be used to assess if convergence has not been achieved.

If all subjects have a monotone missing data pattern (either directly from the study or created by the previous step), the MAR-based multiple imputation with the regression option will be used to impute missing values. This analysis will be performed using SAS PROC MI with the MONOTONE statement and the REGRESSION option with the following specifications:

PROC MI DATA=IN_MCMC NIMPUTE=1 (see note) SEED=234 OUT=OUTPUT;

VAR …; (treatment group, country, class of antidepressant (SNRI or SSRI), and the preceding non-missing values in the order of clinical visits: baseline, Day 2, Day 8, Day 15, Day 22 and Day 28) CLASS…;(treatment group, country, class of antidepressant (SNRI or SSRI)) MONOTONE REGRESSION;

RUN;

Note: NIMPUTE=500 if MCMC was NOT applied at the previous step.

Step 2 –Delta Adjustments

The imputed values for subjects who discontinued will be adjusted by adding δ_p to the imputed values for subjects randomized to placebo and adding δ_A to the imputed values for subjects randomized to an active esketamine arm. Delta-adjusted fully imputed datasets will be generated for different combinations of δ_p and δ_A values as defined below:

- $\delta_p = 0$ and $\delta_A = 0\ to\ \Delta_1^*$ in increments of 1 (active-only adjustment analysis)
- $\delta_p = \delta_A/2$ and $\delta_A = 0\ to\ \Delta_2^*$ in increments of 1 (analysis with a control adjustment that is half of the active adjustment)
- $\delta_p = \delta_A = 0\ to\ \Delta_3^*$ in increments of 1 (all arms adjustment analysis)

Adding positive values results in higher (worse) scores. Δ_1^*, Δ_2^* and Δ_3^* represent the adjustments leading to the 'tipping point', so the smallest delta adjustments values at which conclusions change from favorable to drug (i.e. statistically significant: one-sided p-value ≤ 0.025) to unfavorable (acceptance of the null hypothesis of no treatment difference).

Step 3 –Analysis and Pooling

For each (δ_p,δ_A) combination:

- Same MMRM as described for the primary efficacy analysis will be performed for each set of the adjusted fully imputed datasets;

- Multiple imputation combining rules in PROC MIANALYZE will be applied to the MMRM results from the imputed datasets to produce final inferences.

Between-group comparisons to placebo at Day 28 (e.g., p-values, point estimates for treatment difference) will be displayed graphically for each considered (δ_p, δ_A) combination, up to the 'tipping point' adjustment.

The delta adjustment MI method as described above will be applied to all subjects who discontinued from the DB treatment phase and have missing efficacy scores in this phase. In addition, another version of this method will be applied. In this version, the delta adjustments from Step 2 will be applied to all subjects who discontinued, except those with a discontinuation reason that was not considered related to the study drug, including Lost to Follow-Up, Withdrawal by Subject or Other. A process is in place to obtain clarification on reasons of Withdrawal by Subject and Other to confirm they are not related to the study drug.

Additional Sensitivity Analysis

MMRM will be used for primary efficacy endpoint as described above with the exclusion of 5 subjects from 1 site in the US as a sensitivity analysis due to a potential unblinding issue.

Subgroup Analysis

Forest plots will be provided displaying analysis results for each subgroup listed in Section 2.3. The point estimate of the treatment difference and its 95% confidence interval for each subgroup will be based on an MMRM analysis for the primary endpoint using the appropriate contrast. The model will include baseline MADRS total score as a covariate, and treatment, country, class of antidepressant (SNRI or SSRI), day, subgroup, day-by-treatment interaction, treatment-by subgroup and day-by-treatment-by-subgroup interaction as fixed effects. The terms in the models will be adjusted for the subgroups of country and baseline MADRS total score (⩽/>median). Baseline MADRS total score (as a continuous covariate) will not be included in the model when the dichotomized baseline MADRS total score is included in the model.

5.3. Major Secondary Endpoints

5.3.1. MADRS total score at 24 hours

5.3.1.1. Analysis Methods

The first key efficacy endpoint, change from baseline in MADRS total score at 24 hours post first dose (Day 2), will be analyzed using the same model (MMRM) described above for the MADRS total score at Week 4 in the double-blind treatment phase in Section 5.2.3. The models will include baseline MADRS total score as a covariate, and treatment, country, class of antidepressant (SNRI or SSRI), day (see Table 1), and day-by-treatment interaction as fixed effects.

5.3.2. Sheehan Disability Scale (SDS)

5.3.2.1. 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 scores 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.

5.3.2.2. Analysis Methods

The second key efficacy endpoint, change from baseline in SDS total score at Week 4 in the double-blind treatment phase, will be analyzed using the same model (MMRM) described above for the MADRS total score in Section 5.2.3. Least squares mean changes and arithmetic means for SDS total score over time will be presented graphically. The proportion of subjects who achieve response and remission over time will also be summarized.

Graphs for arithmetic mean change from baseline over time will be provided by country.

5.4. Other Efficacy Variables

5.4.1. Onset of Clinical Response

5.4.1.1. Definition

A subject is defined as having a clinical response if there is at least 50% improvement from baseline in the MADRS total score with onset by Day 2 (in the event that no MADRS was collected on Day2, a MADRS collected on Day 3 could be used) that is maintained to Day 28. Subjects are allowed one excursion (non-response) on Days 8, 15 or 22, however the score must show at least 25% improvement. Subjects who do not meet such criterion, or discontinue during the study before Day 28 for any reason will be considered as non-responders and will be assigned the value of 0 (ie, no).

5.4.1.2. Analysis Methods

The proportion of subjects showing onset of clinical response by Day 2 that is maintained for the duration of the double-blind treatment phase in the esketamine plus oral antidepressant arm will be compared with the oral antidepressant plus intranasal placebo arm using a Cochran-Mantel Haenszel (CMH) chi-square test adjusting for country and class of antidepressant (SNRI or SSRI).

The odds ratio, calculated as the odds of achieving clinical response on esketamine plus oral antidepressant arm divided by the odds of achieving clinical response on oral antidepressant plus intranasal placebo will be provided, along with the 95% confidence interval for the odds ratio.

The proportion of subjects who achieve clinical response will be summarized by class of antidepressant study medication (SNRI or SSRI).

5.4.2. Responders

5.4.2.1. Definition

The percentage change from baseline at Day X is calculated as 100*(MADRS total score at Day X−Baseline MADRS total score)/(Baseline MADRS total score). Negative percent changes in MADRS total score indicate improvement (e.g., percent change ≤ -50% indicates improvement ≥50%).

A subject is defined a responder (yes=1 and no=0) at a given time point if the percent improvement in MADRS total score is ≥50%.

5.4.2.2. Analysis Methods

The proportion of subjects who achieve a response will be summarized by treatment group at each time point during the double-blind treatment and follow-up phases. The proportion of subjects who achieve a response will also be summarized by class of antidepressant study medication (SNRI or SSRI) and by country separately at each time point. Graphs for response rate over time will be provided by country.

The cumulative response rate, defined as the percentage of subjects experiencing at least a given value of percent reduction from baseline to Day 28 in MADRS total score will be presented graphically.

The cumulative distribution function of the time to sustained response will be estimated by the Kaplan-Meier method. Time to sustained response will be summarized (number of sustained responders, number of censored subjects, median, 25^{th} and 75^{th} percentile, if estimable) by treatment group. Sustained response is defined as the first occurrence of response that is maintained through the Day 28 assessment. Subjects are allowed one excursion (non-response) on a subsequent visit prior to Day 28, however the score must show at least 25% improvement. Subjects who discontinue early are not considered to have sustained response.

5.4.3. Improvement of at Least 30% MADRS Total Score

5.4.3.1. Definition

The percentage change from baseline at Day X is calculated as 100*(MADRS total score at Day X−Baseline MADRS total score)/(Baseline MADRS total score). Negative percent changes in MADRS total score indicate improvement (e.g., percent change ≤ -30% indicates improvement ≥30%).

A subject is defined an improvement of ≥30% MADRS total score (yes=1 and no=0) at a given time point if the percent improvement in MADRS total score is ≥30%.

5.4.3.2. Analysis Methods

The analysis will be for overall population and by country separately.

The proportion of subjects who achieve an improvement of ≥30% MADRS total score will be summarized by treatment group at each time point during the double-blind treatment and follow-up phases. Graphs over time will be provided by country. The proportion of subjects who achieve an improvement of ≥30% MADRS total score will also be summarized by class of antidepressant study medication (SNRI or SSRI) and by country separately at each time point.

5.4.4. Remitters

5.4.4.1. Definition

Subjects who have a MADRS total score of ≤12 at a visit will be considered remitters.

5.4.4.2. Analysis Methods

The proportion of subjects who achieve remission will be summarized by treatment group at each time point during the double-blind treatment and follow-up phases. The proportion of subjects who achieve remission will also be summarized by class of antidepressant study medication (SNRI or SSRI) and by country separately at each time point.

The cumulative distribution function of the time to sustained remission will be estimated by the Kaplan-Meier method. Time to sustained remission will be summarized (number of sustained responders, number of censored subjects, median, 25th and 75th percentile, if estimable) by treatment group. Sustained remission is defined as the first occurrence of remission that is maintained through the Day 28 assessment. Subjects are allowed one excursion (MADRS total score >12) on a subsequent visit prior to Day 28. Subjects who discontinue early are not considered to have sustained remission.

5.4.5. CGI-S

5.4.5.1. Definition

The Clinical Global Impression of Severity (CGI-S) 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[4]. 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.

5.4.5.2. Analysis Methods

Descriptive statistics (N, median, minimum, and maximum) of actual values and changes from baseline by treatment group will be provided for observed data for overall population and by

country. Frequency distributions will be provided at each assessment time point during the double-blind and follow-up phases for overall population and by country. Graphs over time will be provided by country.

The change from baseline for CGI-S in the double-blind treatment phase will be analyzed at each time point using the same MMRM as described for primary efficacy endpoint in Section 5.2.3.

Graphs over time will be provided by country.

5.4.6. GAD-7

5.4.6.1. Definition

The GAD-7 (Generalized Anxiety Disorder - 7 Items) 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[10]. 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).

5.4.6.2. Analysis Methods

Descriptive statistics of actual values and changes from baseline by treatment group will be provided for observed case data at all time points in the double-blind treatment and follow-up phases. A frequency distribution will also be provided for GAD-7 severity categories at all time points.

The change from baseline for the GAD-7 total score in the double-blind phase will be analyzed using an ANCOVA model with treatment, country and class of antidepressant (SNRI or SSRI) as factors, and the baseline score as the covariate.

5.4.7. EuroQol Group; 5 Dimension; 5 level (EQ-5D-5L)

5.4.7.1. 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).

5.4.7.2. Analysis Methods

Descriptive statistics of actual values and changes from baseline by treatment group will be provided for the weighted EQ-5D health status index, the EQ-VAS, and the sum score at each time point for the double-blind and follow-up phases.

Individual dimension responses will also be summarized at each visit using a frequency distribution by treatment group for the double-blind and follow-up phases.

5.4.8. SDS in the Follow-up Phase

5.4.8.1. Definition

Definition of SDS is described in Section 5.3.2.1.

5.4.8.2. Analysis Methods

Descriptive statistics of actual values and changes from baseline will be provided at each time point by treatment group for the follow-up phase.

5.4.9. Relapse

5.4.9.1. Definition

During the follow-up phase, relapse (in remitters and responders) is defined as any of the following:

- MADRS total score ≥22 for 2 consecutive assessments separated by 5 to 15 days. The date of the second MADRS assessment will be used for the date of relapse.
- Hospitalization for worsening depression or any other clinically relevant event determined per clinical judgment to be suggestive of a relapse of depressive illness such as suicide attempt, completed suicide, or hospitalization for suicide prevention. If hospitalized for any of these events, the start date of hospitalization will be used for the date of relapse. Otherwise the date of the event will be used if the subject is not hospitalized.
- In case more than one relapse criteria are met, the earlier date will be defined as the date of relapse for this subject.

5.4.9.2. Analysis Methods

For subjects who receive esketamine during the double-blind treatment phase, the proportion of subjects who relapse during the follow-up phase will be summarized separately for subjects who are remitters (MADRS ≤12 at the end of the double-blind phase) and for subjects who are responders but not remitters (≥50% reduction in MADRS total score, but MADRS >12 at the end of the double-blind phase).

In addition, the time to relapse during the follow-up phase will be estimated by the Kaplan-Meier method separately for subjects who are remitters, and subjects who are responders but not remitters. Descriptive statistics (number of relapses, number of censored subjects, and median, 25th, and 75th percentile of time to relapse, if estimable) will be provided.

5.4.10. Six-Item MADRS subscale (MADRS6)

5.4.10.1. Definition

MADRS6[14] is a six-item core depression subscale of MADRS, which contains the following items: apparent sadness, reported sadness, inner tension, lassitude, inability to feel, and pessimistic thoughts.

5.4.10.2. Analysis Methods

Change from baseline in MADRS6 total score at Day 28 in the double-blind treatment phase will be analyzed using the same model (MMRM) described above for the MADRS total score in Section 5.2.3. Least squares mean changes and arithmetic means for MADRS6 total score over time will be presented graphically.

6. SAFETY

All safety summaries for the double-blind treatment phase will be based on the Safety analysis set. Safety summaries for the follow-up phase will be based on the Follow-up analysis set. Unless otherwise specified, the safety data will be summarized for overall population.

6.1. Adverse Events

The verbatim terms used in the CRF by investigators to identify adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). Any AE occurring at or after the initial administration of study agent during the treatment phase through the day of last dose until end of double-blind phase is considered to be treatment emergent.

Treatment-emergent adverse events (TEAEs) are defined as follows for each study phase:

- TEAEs in Double-Blind treatment phase:
  - If AE onset time is not missing:
    a. If intranasal study medication was started on the same day as the oral antidepressant: DB start date/time <= AE onset date and time <= DB end date
    b. If intranasal study medication was not taken or was started after oral antidepressant: DB start date <= AE onset date <= DB end date
  - If AE onset time is missing: DB start date <= AE onset date <= DB 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 assigned based on the rules presented in Section 2.6.

For each adverse event, the number and percentage of subjects who experience at least 1occurrence of the given event will be summarized by treatment group and system organ class and preferred term. The same analysis for the double-blind treatment phase will be provided by country. In addition, comparisons between treatment groups will be provided if appropriate.

Summary tables in the double-blind treatment phase will be provided for:

- TEAEs
- Serious TEAEs
- TEAEs leading to discontinuation of intranasal drug
- TEAEs leading to discontinuation of oral antidepressant
- TEAEs by severity
- TEAEs by relationship to intranasal drug
- AEs by relationship to oral antidepressant
- TEAEs leading to dose reduction

Adverse events occurring in the follow-up phase will be summarized by treatment group separately.

- AEs
- Serious AEs
- AEs leading to discontinuation of oral antidepressant

- AEs leading to termination of study participation
- AEs by severity

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

- Had SAEs
- Had AEs leading to discontinuation of study medication/termination of study participation

In addition, TEAEs will be summarized by severity and relationship to study medication using the preferred term. For the summaries of TEAEs 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 AE reported during the analysis phase by the subject. The proportion of TEAEs occurring on dosing days and the proportion of TEAEs that occur on dosing days with same day resolution will be summarized. Duration and resolution time of severe TEAEs will also be summarized. A listing of subjects who died will be provided. Actual treatment received, cause of death and relationship to study agent (yes/no) will be presented.

Adverse Events of Special Interest

Incidence of other treatment-emergent adverse events of clinical special interest will be summarized.

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

- Drug abuse Potential: 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 synesthetic, 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, Extrasystoles)
- 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, Agitation, Fear, Feeling jittery, Irritability, Nervousness, Panic attack, Tension).
- Events potentially related to suicidality (Completed suicide, Depression suicidal, Intentional overdose, Intentional self-injury, Multiple drug overdose intentional, Poisoning deliberate, Self-injurious behavior, Self-injurious ideation, Suicidal behaviour, Suicidal ideation, Suicide attempt)
- hepatic adverse events (Cholecystitis, Cholelithiasis, Hepatic steatosis, Hepatitis, Non-alcoholic steatohepatitis, Primary biliary cholangitis, Alanine aminotransferase increased, Aspartate aminotransferase increased, Blood bilirubin increased, Gamma-glutamyl transferase increased, Hepatic enzyme increased, Liver function test abnormal, Liver function test increased, Transaminases increased, Urine bilirubin increased, Urobilinogen urine increased)
- events related to renal disorders (Cystitis, Pyelonephritis, Urethritis, Urinary tract infection, Blood creatinine increased, Blood urea increased, Blood urine present, Creatinine renal clearance increased, Protein urine present, Urine analysis abnormal, Urine leukocyte esterase positive, Bladder discomfort, Bladder irritation, Bladder outlet obstruction, Cystitis noninfective, Dysuria, Haematuria, Hypertonic bladder, Lower urinary tract symptoms, Micturition urgency, Nephrolithiasis, Nocturia, Pollakiuria, Polyuria, Proteinuria, Renal colic, Renal failure, Semenuria, Stress urinary incontinence, Ureterolithiasis, Urge incontinence, Urinary bladder polyp, Urinary hesitation, Urinary incontinence, Urinary retention)
- symptoms of dissociation persisting beyond the typical ≤2 hour post esketamine administration (dissociation)
- delirium (Post-injection delirium sedation syndrome, Postoperative delirium, Delirium, Intensive care unit delirium)
- psychosis (Acute psychosis, Affective disorder, Alcoholic psychosis, Bipolar I disorder, Epileptic psychosis, Hysterical psychosis, Mania, Parkinson's disease psychosis, Postictal psychosis, Psychosis postoperative, Psychotic disorder, Psychotic disorder due to a general medical condition, Reactive psychosis, Rebound psychosis, Schizoaffective disorder, Substance-induced psychotic disorder, Transient psychosis)
- mania (Hypomania, Mania)

The number and percentage of subjects taking concomitant medication for dissociation events (preferred term of Dissociation) at any time during the double-blind phase will be provided

6.2. Clinical Laboratory Tests

Descriptive statistics (N, mean, median, minimum, and maximum) for values and changes from baseline will be provided for clinical laboratory tests (hematology, chemistry and urinalysis) at each scheduled time point in the double-blind treatment and follow-up phases. Baseline laboratory result is defined as the last result collected prior to Day 1 predose. This will be used to calculate change for the double-blind treatment phase summary. Changes from baseline for the follow-up phase will also be calculated using the double-blind baseline.

Clinical laboratory tests that meet the criteria for markedly abnormal will be listed by subject for the double-blind treatment phase. The incidence of treatment-emergent markedly abnormal (TEMA) laboratory values that occurred at any time during the double-blind 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 is either missing or within the range given in Attachment 1. If postbaseline 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.

The incidence of subjects with ALT or AST values > 3*upper normal limit (ULN) will be presented for each study phase. Additionally, incidence of hepatic toxicity (Hy's Law[11]) defined as ALT or AST values >3*ULN AND total bilirubin values >2*ULN will be presented for the double-blind treatment and follow-up phases. Similar to the markedly abnormal analysis, only subjects with baseline ALT or AST values ⩽3*ULN (AND baseline total bilirubin values ⩽2*ULN for hepatic toxicity) (or if baseline value is missing) will be eligible for these analyses.

6.3. Vital Signs, Weight and BMI

Descriptive statistics for values and changes from baseline at each scheduled time-point during the double-blind treatment phase will be presented for temperature, systolic blood pressure, diastolic blood pressure, pulse rate, respiratory rate, oxygen saturation, weight, and BMI. In addition, descriptive statistics of pulse rate and blood pressure (systolic and diastolic) values, changes and percent changes from predose will be provided for each dosing day. These summaries will also be provided by hypertension status (history of hypertension recorded in medical history, Yes/No). Frequency distributions of maximum percent change increase from predose and time of maximum percent change increase will also be presented for blood pressure. Descriptive statistics of maximum increase and maximum percent increase from predose will be provided for blood pressure for each dosing day. Note that if the maximum value within a phase occurs at multiple time points, the earliest time point is selected.

The proportion of subjects who have a treatment-emergent abnormality, as defined in Table 7 below, during the double-blind treatment phase will be presented. Both the double-blind baseline and the predose assessment will be used to determine abnormal values. A listing of subjects meeting any of the criteria will also be provided for the double-blind treatment phase.

Table 7: Clinically Important Abnormalities in 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

The proportion of subjects who experienced treatment-emergent markedly elevated blood pressure (systolic BP≥180 or diastolic BP≥110) at any time during the double-blind treatment phase will be summarized by treatment group and hypertension status.

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

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

A separate summary for China population of blood pressure data will be presented for the China population:

- Mean maximum increases from each predose over time
- Frequency of subjects with treatment-emergent abnormal vital signs relative to baseline or predose
- Graph of arithmetic mean
- Frequency of subjects with treatment-emergent markedly elevated blood pressure

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).

The maximum post-baseline value during the double-blind treatment phase will be computed for each ECG parameter using data from both scheduled and unscheduled visits.

The effects on cardiovascular variables will be evaluated by means of descriptive statistics and frequency tabulations. Summary tables for observed values and changes from baseline will be presented by treatment at each scheduled timepoint during the double-blind treatment and follow-up phases.

The frequency of treatment-emergent abnormalities will be tabulated and presented for the double-blind treatment phase. The identification of treatment-emergent abnormal ECG values is based on the post-baseline value (a value occurring after the first study drug administration) being out of range while the baseline value is either missing or within the limits given in Table 8. 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). The average predose value will be used as baseline for the double-blind treatment summary. Abnormal ranges for the HR, PR, QRS and QT intervals are given in Table 8.

Table 8: 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) the incidence of abnormal QTc values and changes from baseline will be summarized by treatment group. Criteria for abnormal corrected QT values and changes from baseline are given in Table 9 and are derived from the ICH E14 Guidance49[5] (the same criteria apply to all QT corrections).

Table 9: 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 for the double-blind and follow-up phases.

The proportion of subjects with treatment emergent abnormalities will be presented for the double-blind treatment phase. A listing of subjects with abnormalities will also be provided.

6.5. Other Safety Parameters

6.5.1. 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 the double-blind baseline for each examination (including the upper respiratory tract/throat) will be listed for the double-blind phase by treatment group.

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

The C-SSRS (Columbia Suicide Severity Rating Scale) 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[9]. 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").

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

A frequency distribution at each scheduled time point by treatment will be provided. Shifts from the baseline visit to the most severe/maximum score during the double-blind treatment and follow-up phases will be summarized by treatment group.

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 visit to the maximum category during the double-blind treatment and follow-up phases will be summarized by treatment group.

6.5.3. Clinician Administered Dissociative States Scale (CADSS)

The CADSS (Clinician Administered Dissociative States Scale) is an instrument for the measurement of present-state dissociative symptoms[1], 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 10.

Table 10: 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 response is missing 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, median, minimum, and maximum) of the total scores and component scores at each time point and visit, changes from predose and proportion of subjects with an increase in CADSS total score from the predose value at any time during the study will be summarized. Mean (SD) CADSS values will be presented graphically for each dosing day.

A separate summary will be presented by country.

6.5.4. 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.

The proportion of subjects with withdrawal symptoms at the end of double-blind treatment therapy or during the follow-up phase will be presented by treatment group. In addition, symptoms at follow-up will be compared to the end of therapy visit and will be summarized using the following categories: new or worsened symptoms, symptoms present and unchanged, no symptoms, and improved.

A separate summary will be presented for the China population.

7. PHARMACOKINETICS

Blood samples for measuring plasma esketamine concentrations will be collected from all subjects at the specified visits as shown in the schedule of events of the protocol.

Data for all subjects who received at least 1 dose of intranasal esketamine and have at least 1 post treatment PK sample will be included in the PK analysis. Esketamine concentrations below the LLOQ of the assay or missing data will be labeled as such in the concentration data listings. No imputation of missing concentration data will be performed, that is, data summaries will be based on the observed data. All subjects and samples excluded from the analysis will be clearly documented.

Descriptive statistics (N, mean, SD, median, range and CV (%)) will be used to summarize esketamine plasma concentrations at each sampling time point by dose.

If feasible, a population PK analysis of plasma concentrations will be performed using nonlinear mixed-effects model (NONMEM) approach. Details will be presented in a separate technical document.

8. BIOMARKER, PHARMACOGENOMIC (DNA), AND EXPRESSION (RNA) EVALUATIONS

Details of the analysis plan for both biomarker and pharmacogenomics analyses will be provided separately based on the availability of sufficient number of subjects.

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-5D5L_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. National Research Council, Panel on Handling Missing Data in Clinical Trials [2010], The prevention and treatment of missing data in clinical trials. Available: http://www.nap.edu/catalog/12955.html.
8. Permutt, T. Sensitivity analysis for missing data in regulatory submissions. Stat Med. 2015; Nov 15 (Epub ahead of print).
9. 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.
10. 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.
11. 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 (Enzyme U/L)	N/A	990
Creatinine [µmol/L]	N/A	265.2
Gamma glutamyl transferase [Enzyme 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 [fraction of 1]	0.3	0.9
Monocytes [fraction of 1]	N/A	0.2
Eosinophils [fraction of 1]	N/A	0.1
Basophils [fraction of 1]	N/A	0.06
Lymphocytes [fraction of 1]	0.1	0.6
Platelet count [$x10^9$/L]	100	600
Red blood cell count [$x10^{12}$/L] -- female	3.0	5.5
-- male	3.0	6.4
White blood cell count [$x10^9$/L]	2.5	15.0
Hy's Law criteria:		
Alanine transaminase (SGPT) [U/L] or Aspartate Aminotransferase (AST)		> 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.