CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

211243Orig1s000

STATISTICAL REVIEW(S)

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA #:	NDA 211-243
Supplement #:	
Drug Name:	SPRATAVO (esketamine) nasal spray: 28 mg, 56 mg, 84 mg
Indication(s):	Treatment of Treatment Resistant Depression
Applicant:	Janssen Pharmaceuticals, Inc.
Date(s):	Submission date: 9/4/2018, PDUFA date: 3/4/2018
Review Priority:	Priority
Biometrics Division:	Division of Biometrics I
Statistical Reviewer:	Andrew N. Potter, PhD
Concurring Reviewers:	Peiling Yang, PhD
	H.M. James Hung, PhD
Medical Division:	Division of Psychiatry Products
Clinical Team:	Jean Kim, MD; Qi Chen, MD MPH
Project Manager:	Hiren Patel, PharmD, RAC

Keywords: mixed models**,** sensitivity analyses, data imputation (except LOCF), multiple comparisons

Table of Contents

1 EXECUTIVE SUMMARY 5

1.1 POSITIVE CONFIRMATORY STUDIES 5
1.2 NEGATIVE CONFIRMATORY STUDIES 5
1.3 CONCLUSIONS AND RECOMMENDATIONS 6

2 INTRODUCTION 7

2.1 OVERVIEW 7
2.2 DATA SOURCES 8

3 STATISTICAL EVALUATION 8

3.1 DATA AND ANALYSIS QUALITY 8
3.2 EVALUATION OF EFFICACY 8
3.2.1 Statistical Review Issues 8
3.2.2 Acute Efficacy Studies 8
3.2.3 Study ESKETINTRD2003 9
3.2.4 Study ESKETINTRD3001 11
3.2.5 Study ESKETINTRD3002 25
3.2.6 Study ESKETINTRD3005 36
3.2.7 Study ESKETINTRD3003 43
3.3 EVALUATION OF SAFETY 56

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS 57

4.1 GENDER, RACE, AGE, AND GEOGRAPHIC REGION 60

5 SUMMARY AND CONCLUSIONS 60

5.1 STATISTICAL ISSUES 60
5.1.1 What esketamine doses (56 mg, 84 mg) shows clinical efficacy? 60
5.1.2 Can the relapse prevention study ESKETINTRD3003 provide confirmatory evidence? 60
5.1.3 Additional Statistics Issues 61
5.2 COLLECTIVE EVIDENCE 61
5.3 CONCLUSIONS AND RECOMMENDATIONS 61

APPENDIX 62

Reference ID: 4392083

LIST OF TABLES

Table 1: List of all studies included in analysis 7
Table 2: MADRS Total Score: Change from Baseline to End Point ANCOVA LOCF Analysis; Double-Blind Phase, Panel A, Periods 1, 2, and Combined (TRD2003: ITT Population) 11
Table 3: Sample Size Re-Estimation Criteria for each Treatment Arm Comparison (TRD3001) 14
Table 4: Interim Analysis Results from DMC Meeting 18 May 2017 16
Table 5: TRD3001 - Number of Patients by Dropout Reason and Treatment Arm 16
Table 6: Demographic Characteristics (TRD3001: FAS) 17
Table 7: MADRS Total Score at Baseline (TRD3001: FAS) 18
Table 8: MADRS Total Score at Four Weeks (TRD3001 FAS) 19
Table 9: Change from Baseline in MADRS Total Score (TRD3001: FAS) 20
Table 10: Subject Disposition (TRD 3002) 27
Table 11: Demographic Characteristics (TRD3002: FAS) 28
Table 12: Baseline MADRS Total Score (TRD3002: FAS) 29
Table 13: MADRS Total Score at 28 Days (TRD3002: FAS) 30
Table 14: Applicant's Analysis of Response Rate at Day 2 (TRD3002: FAS) 30
Table 15: Change from Baseline in MADRS Total Score (TRD3002: FAS) 32
Table 16: Weekly Treatment Effect in MADRS Total Score (TRD3002: FAS) 33
Table 17: Number of Patients per Dose Level (TRD3002: FAS) 35
Table 18: Sample Size Re-Estimation Rules (TRD3005) 37
Table 19: Subject Disposition (TRD3005) 38
Table 20: Demographic Characteristics (TRD3005: FAS) 38
Table 21: Baseline MADRS Total Score (TRD3005: FAS) 39
Table 22: MADRS Total Score at 28 Days (TRD3005: FAS) 40
Table 23: Weekly Treatment Effect in MADRS Total Score (TRD3005: FAS) 41
Table 24: Change from Baseline in MADRS Total Score (TRD3005: FAS) 41
Table 25: Patient Demographic Characteristics (TRD3003: FAS – Remitter Set) 49
Table 26: Patients surviving at least 200 days (TRD3003: FAS - Remitter Set) 54
Table 27: Longitudinal Sub-Model of Joint Model 1 (TRD3003: FAS - Remitter Set) 54
Table 28: Survival Sub-Model of Joint Model 1 (TRD3003: FAS - Remitter Set) 55
Table 29: Longitudinal Sub-Model of Joint Model 2 (TRD3003: FAS - Remitter Set) 55
Table 30: Survival Sub-Model of Joint Model 2 (TRD3003: FAS - Remitter Set) 55

LIST OF FIGURES

Figure 1: Study Schematic for ESKETINTRD2003 10
Figure 2: Study Schematic for ESKETINTRD3001 12
Figure 3: Conditional Power Algorithm for TRD3001 13
Figure 4: TRD 3001 Truncated Fixed Sequence Testing Procedure 15
Figure 5: MADRS Total Score - LS Mean (± SE) Change from Baseline over Time - Mixed Model for Repeated Measures (TRD3001: FAS Population) 20

Reference ID: 4392083

Figure 6: Placebo Subtracted Change from Baseline in MADRS Total Score (TRD3001: FAS) 21
Figure 7: Percentage of Subjects with Specified Change in MADRS Total Score (TRD3001: FAS).....22
Figure 8: Patient Level MADRS Change Score Trajectories (TRD3001: FAS).....23
Figure 9: Esketamine LS Mean Treatment Effect over Study Elapsed Week (TRD3001: FAS)..24
Figure 10: Esketamine One-sided p-value for Treatment Effect over Study Elapsed Week (TRD3001: FAS) 25
Figure 11: Study Schematic for ESKETINTRD3002 26
Figure 12: MADRS Total Score - LS Mean (± SE) Change from Baseline over Time - Mixed Model for Repeated Measures (TRD3002: FAS).....31
Figure 13: MADRS Total Score - LS Mean (± SE) Treatment Effect over Time - Mixed Model for Repeated Measures (TRD3002: FAS) 32
Figure 14: Percentage of Subjects with Specified Change in MADRS Total Score (TRD3002: FAS).....34
Figure 15: Patient Level MADRS Change Score Trajectories (TRD3002: FAS).....35
Figure 16: Sample Esketamine Dose Trajectories (TRD3002: FAS).....36
Figure 17: Study Schematic for ESKETINTRD3005 37
Figure 18: MADRS Total Score - LS Mean (± SE) Change from Baseline over Time - Mixed Model for Repeated Measures (TRD3005: FAS).....40
Figure 19: Percentage of Subjects with Specified Magnitude of Change in MADRS Total Score (TRD3005: FAS) 42
Figure 20: Patient Level MADRS Change Score Trajectories (TRD3005: FAS).....42
Figure 21: Study Schematic for ESKETINDTRD3003.....44
Figure 22: Group Sequential Boundary (TRD3003) 45
Figure 23: Final Log-Rank Test Statistic (TRD3003).....45
Figure 24: Patient Flow Diagram (TRD3003: All Subjects).....48
Figure 25: Patient Disposition (TRD3003: Maintenance Phase - Remitter Set).....48
Figure 26: Time to Relapse (TRD3003: FAS - Remitter Set).....51
Figure 27: MADRS Total Score by Study Day - Optimization Phase Baseline (TRD3003: FAS Remitter Set).....52
Figure 28: MADRS Total Score by Study Day - Optimization Phase Baseline - First Month of Maintenance Phase (TRD3003: FAS Remitter Set) 53
Figure 29: Treatment Arm Average CADSS score Trajectories Stratified by Relapse Status (TRD3003: FAS - Remitter Set).....54
Figure 30: Subgroup Analyses (TRD3001: FAS) 57
Figure 31: Subgroup Analyses (TRD3002: FAS) 58
Figure 32: Subgroup Analyses (TRD3003: FAS - Remitter Set).....59

Reference ID: 4392083

1 EXECUTIVE SUMMARY

Janssen Pharmaceuticals submitted five clinical efficacy studies (ESKETINTRD2003, -TRD3001, -TRD3002, -TRD3003, and -TRD3005) to support a new drug application (NDA) for esketamine with an indication for treatment of treatment-resistant depression (TRD). In addition, the Applicant submitted a study (ESKETINSUI2001) of esketamine in patients with acute suicidal ideation in support of this NDA. Esketamine is the s-enantiomer of ketamine and is approved in several countries as an anesthetic. The Applicant studies esketamine at three doses 28 mg, 56 mg, and 84 mg.

The clinical program consisted of one dose-finding study (TRD2003), three short-term (4 week), parallel group studies (TRD3001, TRD3002, and TRD3005), and a long-term maintenance study (TRD3003). Two (TRD3002 and TRD3003) out of four confirmatory studies provided evidence of esketamine's efficacy. TRD2003 was a two-panel, sequential parallel comparison design (SPCD) that suggested a dose-response where efficacy increases with increasing dose in one panel. However, this dose-response was not confirmed in TRD3001.

1.1 Positive Confirmatory Studies

Studies TRD3002 and TRD3003 both won on their primary statistical hypothesis tests. Study TRD3002 was a four week study comparing flexibly dosed (56 mg to 84 mg esketamine), parallel arm study. The esketamine arm showed a greater decline in depressive symptoms compared to placebo. TRD3003 was a relapse prevention study comparing time to relapse of depressive symptoms between patients stabilized on esketamine (56 mg to 84 mg esketamine) and similar patients who were first stabilized on esketamine then switched to placebo. TRD3003 showed that esketamine increased the time to relapse compared to placebo. However, patients' perception of their treatment assignment may have been influenced by acute side effects (dissociation, sedation, etc.). FDA's exploratory analysis suggested that changes in these side effects were associated with time relapse. However, evidence of esketamine's efficacy remained.

1.2 Negative Confirmatory Studies

Neither study TRD3001 nor TRD3005 provided confirmatory evidence of esketamine's efficacy. TRD3001 compared fixed doses of 56 mg and 84 mg to placebo in adults (18 – 64 years). TRD3001 did not provide evidence that 84 mg esketamine is superior in treating TRD compared to placebo. While 56 mg esketamine arm showed numerical improvement compared to placebo, this dose arm was not formally statistically tested against placebo because the Applicant pre-specified a fixed sequence method to control the type I error. These results contradict the dose-response observed in TRD2003. TRD3005 studied flexibly dosed esketamine (28 mg to 84 mg) in elderly patients (≥ 65 years). In TRD3005, the esketamine and placebo arms did not show any noticeable difference in depressive symptoms except for week 4. The observed treatment effect at week 4 was not statistically significant.

Reference ID: 4392083

1.3 Conclusions and Recommendations

Janssen submitted four confirmatory studies containing mixed evidence of esketamine's efficacy. Two studies (TRD3002 and TRD3003) showed that flexibly dosed esketamine improved depressive symptoms in patients with TRD. However, these two, adequate and well-controlled studies differ from previous antidepressant approvals because previous approvals consist of two short-term, positive studies with the relapse prevention study conducted post-marketing. In addition, there are unanswered questions about the impact of potential changes in patient perception of treatment assignments in TRD3003. Furthermore, the Applicant's submission contained contradictory dose-response findings. We defer to the Division of Psychiatry Products to determine if the positive results in TRD3002 and TRD3003 are sufficient for approval given the concerns about subjects' perception of treatment assignment in TRD3003.

Reference ID: 4392083

2 INTRODUCTION

2.1 Overview

This original NDA contains four phase 3, multi-center, double-blind, placebo-controlled, fixed and flexible dose, parallel-group studies designed to evaluate the efficacy and safety of esketamine in adults 18+. Esketamine was studied in doses 28 mg, 56 mg, and 84 mg. The dosing frequency varied from every third day during the first four weeks of treatment to every other week to maintain efficacy.

The original protocol was reviewed under IND 114,345.

Table 1: List of all studies included in analysis

	Phase and Design*	Treatment Period	# of Subjects per Arm	Study Population
ESKETINTRD 2003	*Phase 2 – MC, R, DB, SPCD, trial*	*8 days*	*Esketamine 28mg / P1: 11, P2: 16* *Esketamine 56mg / P1: 11, P2: 20* *Esketamine 84mg / P1: 12, P2: 17* *Placebo / P1: 33, P2: 10*	*Adults ages 20-64 with TRD*
ESKETINTRD 3001	*Phase 3 - MC, R, DB, PG, AC trial*	*4 weeks*	*Esketamine 56mg / 115* *Esketamine 84mg / 114* *Placebo / 113*	*Adults ages 18-65 with TRD*
ESKETINTRD 3002	*Phase 3 - MC, R, DB, PG, AC trial*	*4 weeks*	*Esketamine / 114* *Placebo / 109*	*Adults ages 18-65 with TRD*
ESKETINTRD 3003	*Phase 3 - MC, RW, DB, PG, AC trial*	*Event driven*	*Esketamine** / 114* *Placebo / 109*	*Adults ages 18-65 with TRD*
ESKETINTRD 3005	*Phase 3 - MC, R, DB, PG, AC trial*	*4 weeks*	*Esketamine / 72* *Placebo / 65*	*Adults ages 65+ with TRD*

* MC: multi-center, R: randomized, DB: double-blind, PG: parallel group, AC: active controlled, RW: randomized withdrawal
** Sample size reported for the complete remission set used for the primary efficacy analysis.

Reviewer's Note: *The Applicant submitted a sixth study, ESKETINSUI2001, as an additional supportive study. However, this study was conducted in patients with acute suicidal ideation and used a primary endpoint at 4 hours post-dose instead of 8 or 28 days after initial dosing. Because of a different study population and indication, SUI2001 is not reviewed in this application.*

Reference ID: 4392083

2.2 Data Sources

The following data sources were considered in this review:
Study ESKETINTRD 3001: adsl, admadrs, zr, adbdc
Study ESKETINTRD 3002: adsl, admadrs, adex
Study ESKETINTRD 3003: adsl, admadrs, adtte, ce, adcadss
Study ESKETINTRD 3005: adsl, admadrs

All studies are referred to by TRD 300X. The placebo + oral antidepressant arm will be referred to as placebo or placebo + oral antidepressant.

The electronic location of the submission is: \\cdsesub1\evsprod\NDA211243\0001.

3 STATISTICAL EVALUATION

3.1 Data and Analysis Quality

The Applicant submitted all necessary analysis datasets and SAS programs. This reviewer found the datasets acceptable. For TRD3001, the Applicant submitted the requested DMC interim analysis minutes for the closed session and the interim analysis statistics report.

3.2 Evaluation of Efficacy

In all studies supporting this NDA, treatment resistant depression (TRD) was defined as failure of two adequate previous anti-depressant treatments. The Applicant required patients to fail a treatment during the screening period of the study.

3.2.1 Statistical Review Issues

- What esketamine doses (56 mg, 84 mg) shows clinical efficacy?
- Can a relapse prevention trial provide confirmatory evidence when the treatment has acute side-effects? Acute side-effects include dissociative and sedative effects with the first hour after dose administration.
- What statistical evidence can be extracted from TRD 3001 after statistical testing stopped at 84 mg dose?
- Does the esketamine treatment effect improve over time?

3.2.2 Acute Efficacy Studies

Studies TRD3001, TRD3002, and TRD3005 shared multiple design features:

- All three studies were four week, double-blind, randomized studies comparing esketamine plus a newly initiated oral antidepressant (OAD) to a newly initiated oral anti-depressant. Patients started on duloxetine (SNRI), venlafaxine XR (SNRI), sertraline (SSRI), or escitalopram (SSRI) depending on the patients' prior drug exposure and physician choice. Randomization was stratified by new OAD class.

- During the four week double phase, patients took esketamine every three days.

Reference ID: 4392083

- The primary efficacy endpoint was the MADRS total score. The MADRS rating scale measures depressive symptoms on 10 items. The item level scores range from 0 to 6 points. The total score ranges from 0 to 60.

- MADRS total score was measured 1 day (except for TRD3005), 7 days, 14 days, and 28 days after first dose of esketamine.

- For the primary endpoint of change in MADRS score from baseline to day 28, the primary analysis was a MMRM analysis. The model included fixed effect terms for treatment arm, categorical visit, treatment-by-visit interaction, baseline MADRS score, oral antidepressant class, and pooled region. An unstructured covariance matrix was pre-specified. The model was fit using restricted maximum likelihood (REML). If the unstructured covariance matrix model did not converge, the Applicant proposed to try the following covariance structures: heterogenous Toeplitz, Toeplitz, heterogenous auto-regressive, and auto-regressive. The denominator degrees of freedom were estimated using Kenward-Roger approximation. The esketamine treatment effect for both treatment arms was the least squares (LS) mean difference at day 28 from this MMRM model.

- All efficacy analyses were conducted using the Full Analysis Set (FAS). The FAS population consisted of all randomized subjects who received at least one dose of esketamine and at least one dose of oral antidepressant.

Reviewer's Note: *The Clinical Reviewer, Dr. Jean Kim, raised concerns that the concomitant starting of esketamine and a new OAD prevents the study of the effect of esketamine alone. Dr. Kim also raises concerns that the label will require a change in OAD with starting esketamine.*

3.2.3 Study ESKETINTRD2003

TRD2003 was a two panel, fixed dose, placebo-controlled, double-blind, multicenter, sequential parallel comparison design (SPCD). Panel A enrolled 67 patients comparing 28 mg, 56 mg, and 84 mg esketamine to placebo (see Figure 1); Panel B enrolled 41 patients comparing 14 mg and 56 mg esketamine to placebo. The primary efficacy endpoint was MADRS change from baseline to 8 day after esketamine dosing. Panel A was conducted in the United States and Belgium. Panel B was conducted in Japan. Only Panel A is reviewed.

In Panel A - Period 1, patients were randomized at a 3:1:1:1 ratio to placebo and 28 mg, 56 mg, and 84 mg esketamine. The Applicant assumed that the mean MADRS change from baseline was 9 points with a standard deviation of 10 and an alpha level of 0.1. If a placebo patient did not respond by day 8 (Quick Inventory of Depressive Symptomatology (QIDS) score ≥ 11), they were re-randomized to placebo, 28 mg, 56 mg, and 84 mg at a 1:1:1:1 ratio (Period 2). The Applicant further assumed Period 2 treatment effect is a mixture of 7 points (40% of these non-responders with a moderate QIDS score) and 9 points (55% of these non-responders with a several QIDS score). In Period 1, pairwise differences were estimated using an ANCOVA model with factors of country and treatment with a continuous covariate of baseline MADRS score. In Period 2, pairwise differences were estimated using an ANCOVA model with factors of country and treatment with a continuous covariate of baseline (end of Period 1) MADRS score

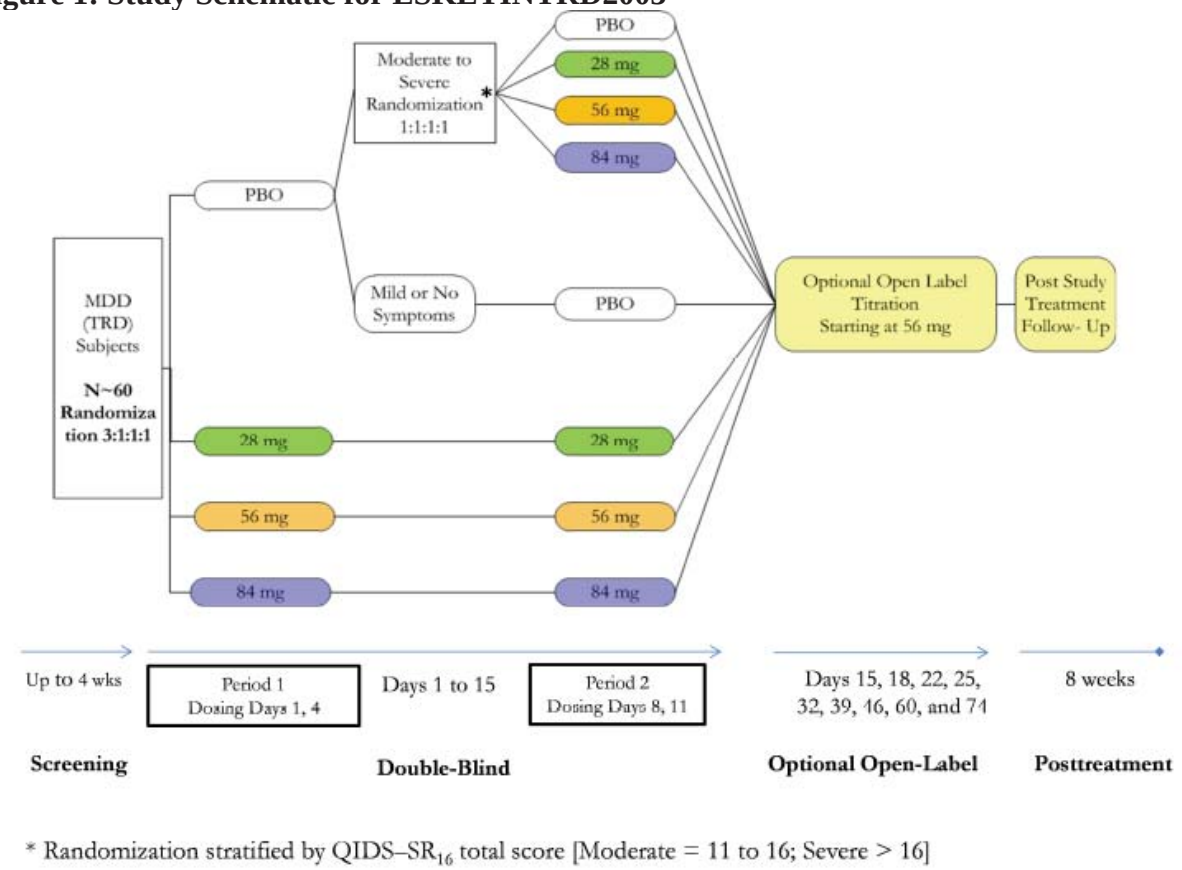
Reference ID: 4392083

and baseline QIDS score. Results from Periods 1 and 2 were combined using two different weighted combination tests:

- [redacted] (b) (4) is the first period test statistic, and Z_2 is the first period test statistic
- Test statistic from the Applicant's technical paper

Reviewer's Note: *In the Statistical Analysis Plan (SAP), the Applicant did not specify any method to control type I error over the primary endpoint family (F1) all pairwise comparisons between esketamine and placebo. In addition, the Applicant did not pre-specify which weighted combination test would be used for statistical inference. Therefore, the actual type I error spent on F1 is not known.*

Figure 1: Study Schematic for ESKETINTRD2003



Source: Applicant's clinical study report.

In Table 2, results from Panel A (Period 1 and Period 2) suggest that the efficacy of esketamine increases with increasing esketamine dose. Supporting this trend are three statistically significant (at a nominal significance level of 0.1), pairwise differences in mean MADRS total score in all esketamine dose groups versus placebo for Period 1; Period 2; and Periods 1 and 2 combined on ANCOVA (LOCF) analysis. The only exception is the pairwise comparison between 28 mg esketamine and placebo in Period 2 (unadjusted p-value = 0.152).

Reference ID: 4392083

Table 2: MADRS Total Score: Change from Baseline to End Point ANCOVA LOCF Analysis; Double-Blind Phase, Panel A, Periods 1, 2, and Combined (TRD2003: ITT Population)

Period		*Placebo*	*Esketamine 28 mg*	*Esketamine 56 mg*	*Esketamine 84 mg*
1	N	33	11	11	12
	LS Mean (SE)	-4.9(1.74)	-9.8 (2.72)	-12.4 (2.66)	-15.3 (2.56)
	Mean diff Placebo (SE)	--	-5.0 (2.99)	-7.6 (2.90)	-10.5 (2.79)
	One-sided p-value < 0.025	--	0.051	0.006	<0.001
2	N	6	8	9	5
	LS Mean (SE)	-4.5 (2.92)	-7.6 (2.49)	-8.9 (2.51)	-11.4 (2.68)
	Mean diff Placebo (SE)	--	-3.1 (2.99)	-4.4 (3.06)	-6.9 (3.41)
	One-sided p-value < 0.025	--	0.152	0.083	0.028
Combined	Mean diff Placebo (SE)	--	-4.2 (2.1)	-6.3 (2.1)	-9.0 (2.1)
	90% CI	--	-7.7 to -0.8	-9.7 to -2.9	-12.5 to -5.5
	One-sided p-value < 0.025	--	0.021	0.001	<0.001

Source: Applicant's Clinical Study Report.

Reviewer's Note: *This was a dose-finding study. The Applicant conducted multiple statistical tests (3 pairwise comparisons in the primary family) using two different weighted combination tests without pre-specifying any order of testing. In addition, Study 2003 may not have met the prespecified assumptions for the validity of either of the SPCD combination tests conducted. However, the Period 1 results suggest increasing antidepressant efficacy with increasing esketamine dose. This finding does not depend on meeting the SPCD assumptions for combining results between periods.*

3.2.4 Study ESKETINTRD3001

3.2.4.1 Study Design and Endpoints

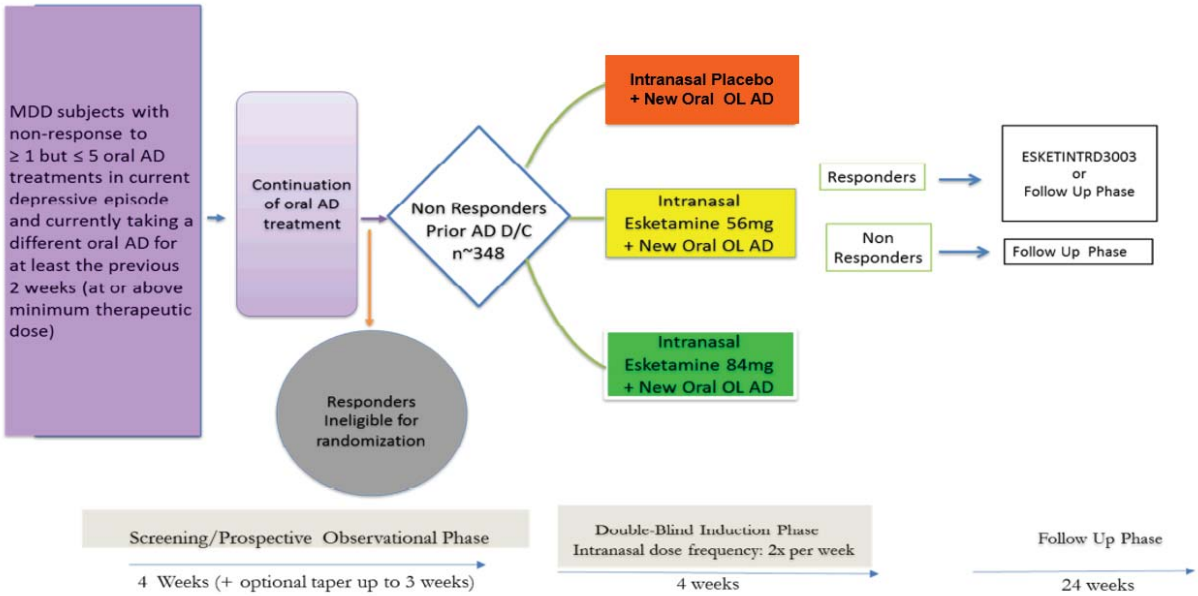
TRD3001 was a fixed dose study comparing 56 mg and 84 mg esketamine + OAD to placebo + OAD. The study was conducted at 96 sites worldwide with 42 sites in the United States. Subjects were randomized at a 1:1:1 ratio to either 56 mg esketamine, 84 mg esketamine, or

Reference ID: 4392083

placebo. The Applicant planned to originally enroll 234 patients with a sample size re-estimation after 50% of patients were enrolled. The maximum sample size was 348. Key secondary endpoints were:

- Onset of clinical response by Day 2
- Functioning and associated disability measured by Sheehan Disability Scale (SDS)
- Subject reported depressive symptoms measured by PHQ-9

Figure 2: Study Schematic for ESKETINTRD3001



Source: Adapted from Applicant's Analysis Data Reviewer's Guide for TRD3001.

3.2.4.2 **Statistical Methodologies**

<u>Applicant's Methods</u>

Study TRD3001 was designed to test two dose arms (56 mg, 84 mg every three days) vs. placebo controlling for multiple comparisons using a fixed sequence procedure, see Figure 4. The one-sided, family-wise type I error rate (FWER) was controlled at 2.5%. The Applicant designed TRD3001 to have 90% power at its maximum sample size of 348 with a one-sided alpha level of 1.25% and 25% dropout rate. The Applicant further assumed that the treatment difference between placebo and each esketamine arm was 6.5 point change on the MADRS at week 4 with a standard deviation of 12. These assumptions were based on Phase 2 results from 67 patients. The Applicant also estimated the minimum sample size of 234 by changing the assumed treatment effect of 8 point change in MADRS.

<u>Reviewer's Note:</u> *In a Phase 2, 67 patient, dose-ranging trial TRD2003, the Applicant studied esketamine at doses 28 mg, 56 mg, and 84 mg compared to placebo (not required to switch to a novel oral antidepressant). At 8 days post-dose, the placebo adjusted change from baseline in MADRS score was -4.2 (-7.67, -0.79), -6.03 (-9.71, -2.88), and -9.0 (-12.53, -5.52) for 28 mg, 56 mg, and 84 mg esketamine respectively. The observed treatment effect for TRD3001 was approximately -4.0 points on the MADRS for both 56 mg and 84 mg esketamine. TRD3002 and*

Reference ID: 4392083

TRD3005 showed similar magnitudes of treatment difference from placebo. If TRD3001 were designed using the Phase 3 treatment difference from placebo of -4.0 and SD = 12, the total maximum sample size would rise to 675 patients.

The Applicant conducted an interim analysis to re-estimate sample size and test for futility after 121 patients were enrolled. The Applicant used an unblinded, conditional power approach to re-estimate the sample size and determine futility. The estimated treatment effects were assumed to be the true effect for the conditional power calculations. An independent Data Monitoring Committee (DMC) estimated conditional power using the algorithm outlined in Figure 3 for both 56 mg esketamine and 84 mg esketamine arms. Using these estimated conditional powers, the DMC used the criteria in Table 3.

Figure 3: Conditional Power Algorithm for TRD3001

1. At the time of the interim analysis, conditional power (CP) will be calculated assuming that the estimated treatment effect observed at the interim analysis is the true effect using the formula [redacted] (b) (4) where,

Φ: cumulative distribution function of the standard normal distribution

z: Observed normalized test statistic for esketamine dose versus active comparator comparison

$z_{\alpha/2}$: critical value of z at $\alpha/2$ = [redacted] (b) (4)

t: IA information fraction is given by the following formula

[redacted] (b) (4)

* Planned total number of subjects: 234

Source: Applicant's Analysis Data Reviewer's Guide for TRD3001.

Reference ID: 4392083

Table 3: Sample Size Re-Estimation Criteria for each Treatment Arm Comparison (TRD3001)

Scenario	Conditional Power (%)		Re-estimated total sample size
	Comparison 1	Comparison 2	
1	$CP_1<10$	$CP_2<10$	Stop study
2a	$CP_1<10$	$10\leq CP_2\leq 30$	234
2b	$10\leq CP_1\leq 30$	$CP_2<10$	234
2c	$10\leq CP_1\leq 30$	$10\leq CP_2\leq 30$	234
3a	$CP_1\leq 30$	$30<CP_2<50$	348
3b	$30<CP_1<50$	$CP_2\leq 30$	348
3c	$30<CP_1<50$	$30<CP_2<50$	348
4a	$CP_1\leq 30$	$50\leq CP_2<80$	291
4b	$50\leq CP_1<80$	$CP_2\leq 30$	291
4c	$50\leq CP_1<80$	$50\leq CP_2<80$	291
5a	$CP_1\leq 30$	$CP_2\geq 80$	234
5b	$CP_1\geq 80$	$CP_2\leq 30$	234
5c	$CP_1\geq 80$	$CP_2\geq 80$	234
6a	$30<CP_1<50$	$50\leq CP_2$	348
6b	$50\leq CP_1$	$30<CP_2<50$	348
7a	$50\leq CP_1<80$	$CP_2\geq 80$	291
7b	$CP_1\geq 80$	$50\leq CP_2<80$	291

Source: Applicant's Analysis Data Reviewer's Guide for TRD3001.

For the primary analysis, the Applicant calculated treatment effect point estimates, interval estimates, and p-values using a weighted combination of results from two separate MMRM analyses described in Section 3.2.2. A weighted combination test used the test statistic

(b) (4)

wher (b) (4) denote the first and second stage z-values corresponding to the stage-wise p-values $p_1$and p_2. Data for each stage is analyzed separately using MMRM. The median, unbiased, combined treatment effect, $\hat{\delta}$, is estimated by

(b) (4)

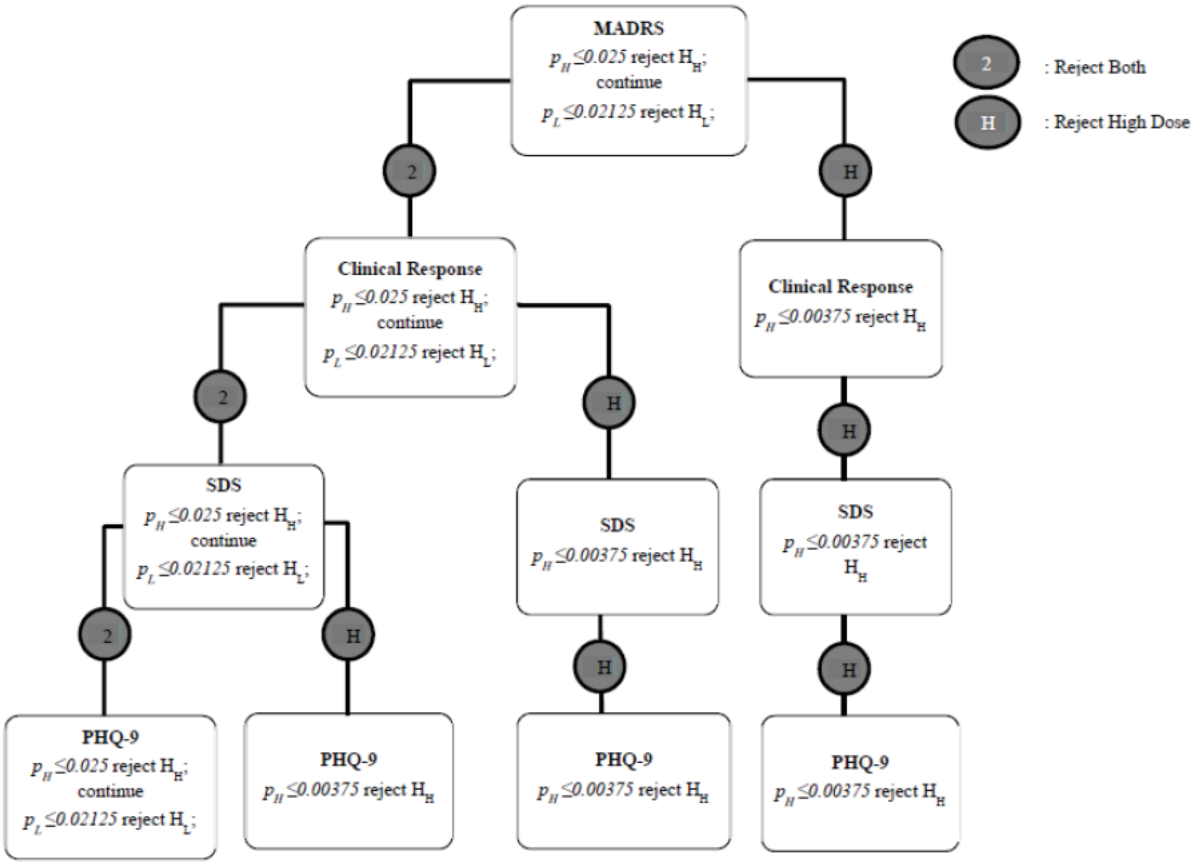
where $\hat{\delta}_1$ and $\hat{\delta}_2$ are the stage-wise treatment effects, $\mathrm{SE}(\hat{\delta}_1)$ and $\mathrm{SE}(\hat{\delta}_2)$ are the respective standard errors, and (b) (4) are pre-specified weights. A corresponding (b) (4) two-sided confidence interval is

Reference ID: 4392083

(b) (4)

where z_α is the (b) (4) -th percentile of the standard normal distribution. The stage-wise estimates are the least square means (LSmean) and associated standard errors.

Figure 4: TRD 3001 Truncated Fixed Sequence Testing Procedure



Source: Applicant's SAP for TRD3001.

3.2.4.3 Interim Analysis Results

The preplanned interim analysis was conducted after 121 patients were randomized and discussed at the closed DMC session on 18 May 2017. The independent statistician included 107 patients in the interim analysis. For the 56 mg arm, the conditional power was 17.78%, see Table 4. For the 84 mg arm, the conditional power was 0.04%. The DMC recommended that the study continue with the minimum sample size of 234 patients. When the randomization vendor notified the Applicant that 234 patients were randomized, additional patients had been enrolled in the screening period. The Applicant decided that they had an ethical obligation to offer enrollment in TRD3001 if a patient met entry criteria. The final sample size was 346 with

Reference ID: 4392083

342 in the FAS. The final sample 346 was 2 fewer patients than the planned maximum sample size of 348.

Table 4: Interim Analysis Results from DMC Meeting 18 May 2017

	Intranasal Esk 56 mg + Oral AD (N=37)	Intranasal Esk 84 mg + Oral AD (N=42)	Oral AD + Intranasal Placebo (N=42)
Baseline(IND)			
N	37	42	42
Mean (SD)	36.6 (4.62)	37.5 (5.82)	36.6 (7.34)
Median (Range)	37.0 (27; 47)	37.5 (26; 49)	36.0 (18; 53)
Day 28			
N	34	35	38
Mean (SD)	15.1 (12.27)	19.2 (13.16)	19.6 (13.69)
Median (Range)	16.0 (0; 39)	18.0 (1; 41)	20.5 (0; 48)
Change from Baseline to Day 28			
N	34	35	38
Mean (SD)	-21.5 (12.52)	-18.5 (13.76)	-17.5 (17.60)
Median (Range)	-22.5 (-41; 6)	-20.0 (-41; 12)	-16.5 (-51; 21)
1-sided P-value (minus Placebo)(a)	0.126	0.525	
Diff. of LS Means (SE)	-3.6 (3.12)	0.2 (3.08)	
95% CI (b)	(-9.77;2.59)	(-5.90;6.29)	
Conditional power assuming current trend (%)	17.78	0.04	
Recommended decision (c)			Continue Study
Recommended total sample size (c)			234

a) Based on a mixed-effects model for repeated measures (MMRM) with treatment, class of oral antidepressant (SSRI, SNRI), region, day, and day-by-treatment interaction as fixed effects, baseline value as a covariate, and a random subject effect.
b) The confidence intervals are unadjusted for multiplicity.
c) Based on conditional power (CP), the total sample size is re-estimated based on the rules presented on the next page.
Note: Negative change in score indicates improvement.

Source: Applicant's independent statistics report for TRD3001 interim analysis.

3.2.4.4 Protocol Amendments

On 16 May 2016, the Applicant made a substantial amendment to the protocol. The Applicant changed the inclusion criteria to allow patients to enter the screening period having failed at least one adequate antidepressant therapy instead of at least two adequate antidepressant therapies. No amendments affected the statistical plan.

3.2.4.5 Patient Disposition, Demographic and Baseline Characteristics

Of the 342 randomized subjects, 113 subjects were randomized to placebo, 117 subjects were randomized to 56 mg esketamine, and 116 subjects were randomized to 84 mg esketamine. Eight percent of subjects dropped out of the double-blind (DB) phase. The 84 mg dose arm had a higher dropout rate of 15%. Additional details of patient disposition are presented in Table 5.

Table 5: TRD3001 - Number of Patients by Dropout Reason and Treatment Arm

	Esketamine

Reference ID: 4392083

	Placebo – N(%)	56 mg – N(%)	84 mg – N(%)	Overall – N(%)
Randomized	113	117	116	346
Full Analysis Set	113	115	114	342
Completed 28-day treatment	107 (94.69)	111 (96.52)	97 (85.09)	315 (92.11)
Dropout – Adverse Event	2 (1.78)	1 (0.87)	7 (6.14)	10 (2.92)
Dropout – Lack of Efficacy	0 (0)	1 (0.87)	1 (0.88)	2 (0.58)
Dropout – Other Reasons	4 (3.54)	2 (1.74)	9 (7.89)	15 (4.39)

Source: Reviewer.

In the FAS, the patients were majority female (70%). In addition, the FAS was predominantly white (77%). Patients were an average age of 46 years. Patients were overweight with a mean BMI of 27.4 kg/m^2. Approximately 40% of patients were from the United States. Demographic variables were well balanced across all treatment arms. Patients in the 84 mg dose arm were more likely to have failed at least three previous oral antidepressant therapies (48%) compared to 30% in the 56 mg dose arm and 41% in the placebo arm. See Table 6 for additional demographic characteristics.

Table 6: Demographic Characteristics (TRD3001: FAS)

		Esketamine			
Characteristic	**Statistics**	**56 mg N = 115**	**84 mg N = 114**	**Placebo N = 113**	**Overall N =342**
Age (years)	Mean	46.43	45.73	46.78	46.31
	SD	11.18	11.10	11.36	11.19
	Median	48.00	47.00	47.00	47.00
	Min, Max	22, 64	18, 64	18, 64	18, 64
Sex, n (%)					
Male		34 (30)	35 (31)	32 (28)	101 (30)
Female		81 (70)	79 (69)	81 (72)	241 (70)
Race, n (%)					
American Indian or Alaska Native		0	1 (1)	0	1 (0)
Asian		2 (2)	1 (1)	2 (2)	5 (1)
Black or African American		7 (6)	7 (6)	5 (4)	19 (6)
Multiple		0	0	1 (1)	1 (0)
Not Reported		7 (6)	9 (8)	9 (8)	25 (7)
Other		8 (7)	11 (10)	10 (9)	29 (8)
White		91 (79)	85 (75)	86 (76)	262 (77)
Ethnicity, n (%)					
Hispanic or Latino		33 (29)	27 (24)	31 (27)	91 (27)

Reference ID: 4392083

Not Hispanic or Latino		82 (71)	87 (76)	82 (73)	251 (73)
Weight (kg)	Mean	80.98	80.55	83.33	81.61
	SD	20.55	20.16	21.92	20.86
	Median	77.60	76.95	79.50	77.60
	Min, Max	44.1, 156.9	50.4, 149.2	49.0, 170.9	44.1, 170.9
Height (cm)	Mean	167.56	168.09	168.60	168.08
	SD	9.74	9.85	9.98	9.8
	Median	166	168	168	167
	Min, Max	149, 198	145.5, 193	141, 200	141, 200
BMI (kg/m²)	Mean	28.76	28.37	29.20	28.77
	SD	6.70	5.86	6.69	6.42
	Median	27.30	28.15	27.50	27.40
	Min, Max	18.00, 55.90	17.00, 49.90	19.00, 49.70	17.00, 55.90
Country, n(%)					
Belgium		8 (7)	9 (8)	12 (11)	29 (8)
Brazil		20 (17)	19 (17)	18 (16)	57 (17)
Canada		7 (6)	7 (6)	6 (5)	20 (6)
Estonia		3 (3)	4 (4)	3 (3)	10 (3)
France		11 (10)	10 (9)	10 (9)	31 (9)
Hungary		3 (3)	1 (1)	1 (1)	5 (1)
Mexico		14 (12)	16 (14)	15 (13)	45 (13)
Slovakia		4 (3)	3 (1)	3 (3)	10 (3)
United States		45 (39)	45 (39)	45 (40)	135 (39)
Novel Oral Antidepressant, n(%)					
Duloxetine		49 (43)	43 (38)	44 (39)	136 (40)
Escitalopram		26 (23)	23 (20)	24 (20)	73 (21)
Sertraline		24 (21)	24 (21)	25 (22)	73 (21)
Venlafaxine XR		16 (14)	24 (21)	20 (18)	60 (18)
Failed Oral Antidepressant Therapies, n(%)					
1		14 (12)	12 (11)	7 (6)	33 (10)
2		67 (58)	47 (41)	60 (53)	174 (51)
3 or more		34 (30)	55 (48)	46 (41)	135 (39)

Source: Reviewer.

At baseline, subjects had an average, baseline MADRS total score of 37.45 in the placebo arm, 37.42 in the 56 mg arm, and 37.84 in the 84 mg arm.

Table 7: MADRS Total Score at Baseline (TRD3001: FAS)

			Esketamine	
Scale	**Statistics**	**Placebo (N=113)**	**56 mg (N=115)**	**84 mg (N=114)**
MADRS	Mean	37.45	37.42	37.84
	SD	6.16	4.76	5.58
	Median	37	37	37.5

Reference ID: 4392083

Min, Max	18, 53	27, 50	25, 51

Source: Reviewer.

3.2.4.6 Results and Conclusions

Applicant's Results

The primary efficacy endpoint of change from baseline to week 4 in MADRS total score in the 84 mg dose arm did not reach statistical significance (one-sided p = 0.0353). Statistical testing stopped after the 84 mg dose arm failed. In the 84 mg dose arm, mean MADRS total score declined by 3.34 (95% CI[1]: -0.27, 6.95) points more than the placebo arm. The 56 mg/day dose arm had a larger difference of 4.12 (95% CI: 0.58, 7.66) in decline. Detailed results are found in Table 8. All p-values in Table 8 are one-sided p-values. Combined p-values were compared to an alpha level of 0.025 for the 84 mg dose and 0.02125 for the 56 mg dose if the 84 mg dose were statistically significant. Throughout this Section, negative change indicates improvement.

At the end of stage 1, the 84 mg dose arm showed little difference from placebo (-0.29 points greater change from baseline). At the end of stage 2, the 84 mg dose arm average MADRS score declined by -5.69 points more than placebo. Patient baseline demographic characteristics and medical history did not differ between stages.

Table 8: MADRS Total Score at Four Weeks (TRD3001 FAS)

Endpoint	Dose	Placebo		Esketamine		LS Mean Difference (SE/ 95% CI)	One-sided p-value
		N	LS Mean (SE/ 95% CI)	N	LS Mean (SE/ 95% CI)		
Stage 1	56 mg	42	-16.83 (2.09)	37	-20.25 (2.22)	-3.42 (3.03)	0.1304
	84 mg			42	-17.12 (2.18)	-0.29 (3.00)	0.4617
Stage 2	56 mg	71	-13.39 (1.61)	78	-18.02 (1.53)	-4.63 (2.21)	0.0181
	84 mg			72	-19.07 (1.68)	-5.68 (2.30)	0.0098
Combined	56 mg	113	-14.89 (-17.41, -12.37)	115	-18.93 (-21.44, -16.42)	-4.12 (-7.66, -0.58)	0.0114
	84 mg			114	-18.22 (-20.85, -15.59)	-3.34 (-6.95, 0.27)	0.0353

Source: Reviewer.

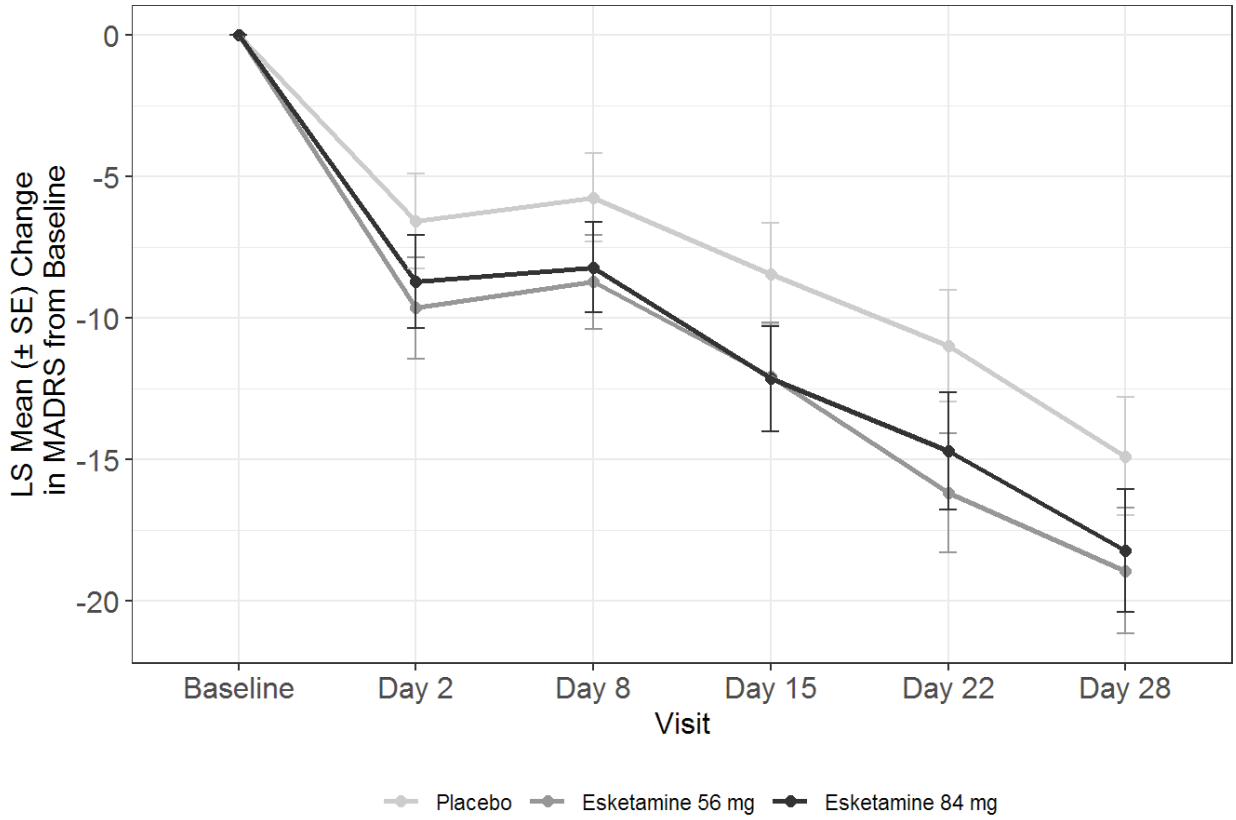
Over the four-week study period, the mean MADRS total score generally declined in all dose arms, see Figure 5 and Table 9. At 24 hours after the first dose, all dose arms showed a decline from baseline. MADRS score showed little change over the next week. From day 8 to day 28, all dose arms showed an approximately linear decline in MADRS scores. The treatment effect relative to placebo for both esketamine doses ranged between -2 and -5 points throughout the 4 week study, see Figure 6. While there may be a time trend towards increasing treatment difference from placebo, there is no statistical evidence for this trend. In TRD3001, the 84 mg dose arm showed similar or weaker efficacy compared to the 56 mg dose arm (84 mg: -3.34 vs. 56 mg: -4.12). TRD3001 suggests that esketamine's efficacy does not increase with dose. This result contradicts the dose response observed in TRD2003.

[1] Confidence interval is abbreviated CI.

Reference ID: 4392083

Reviewer's Note: *In the 84 mg dose arm, this Reviewer did not find any obvious reasons for difference between stage 1 LS mean difference (-0.29) and stage 2 LS mean difference (-5.68).*

Figure 5: MADRS Total Score - LS Mean (± SE) Change from Baseline over Time - Mixed Model for Repeated Measures (TRD3001: FAS Population)



Source: Reviewer

Table 9: Change from Baseline in MADRS Total Score (TRD3001: FAS)

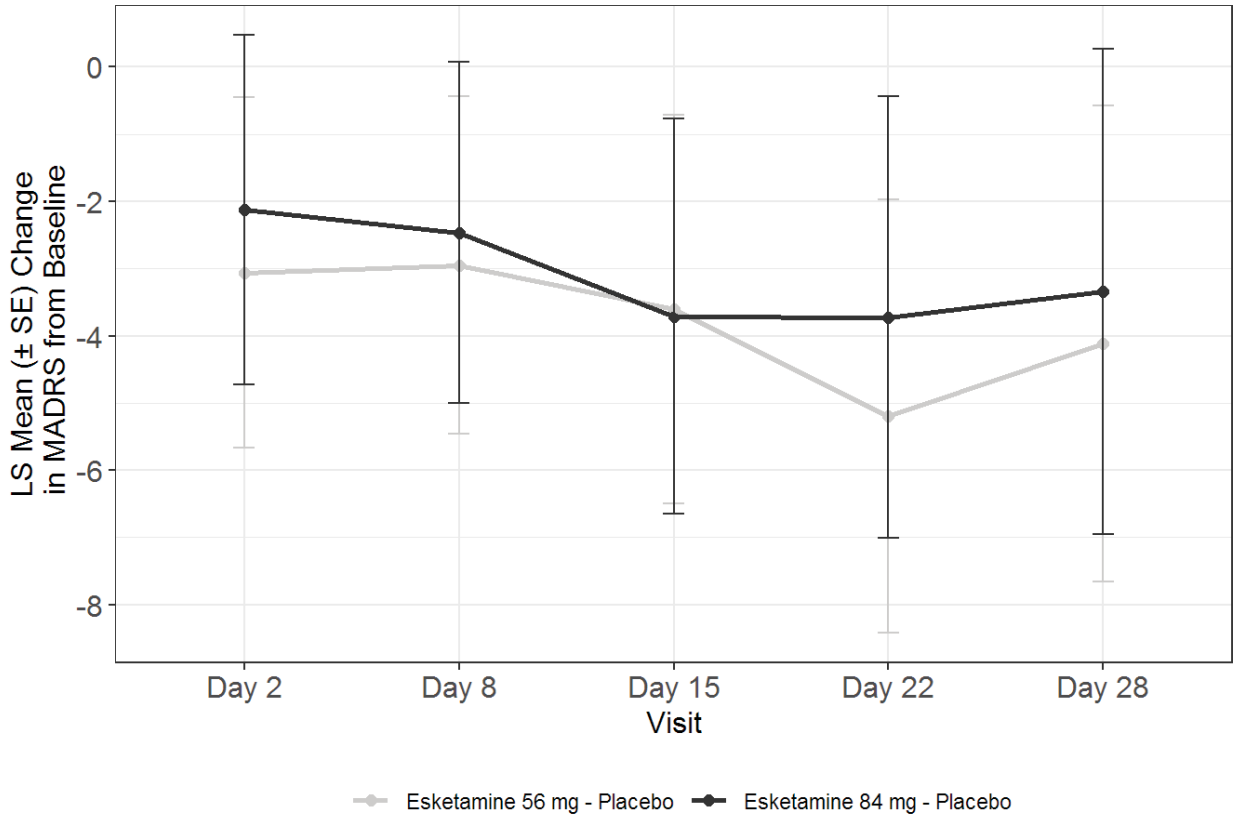
Visit	*Treatment Arm*	*Change from Baseline*	*Lower 95% CI*	*Upper 95% CI*
Day 2	Placebo	-6.58	-8.46	-4.71
	Esketamine 56 mg	-9.64	-11.49	-7.78
	Esketamine 84 mg	-8.71	-10.58	-6.84
Day 8	Placebo	-5.74	-7.54	-3.94
	Esketamine 56 mg	-8.72	-10.52	-6.92
	Esketamine 84 mg	-8.20	-10.05	-6.35
Day 15	Placebo	-8.44	-10.51	-6.38
	Esketamine 56 mg	-12.08	-14.14	-10.02
	Esketamine 84 mg	-12.14	-14.29	-10.00
Day 22	Placebo	-10.97	-13.28	-8.67
	Esketamine 56 mg	-16.17	-18.46	-13.88
	Esketamine 84 mg	-14.70	-17.10	-12.30
Day 28	Placebo	-14.89	-17.41	-12.37

Reference ID: 4392083

	Esketamine 56 mg	-18.93	-21.44	-16.42
	Esketamine 84 mg	-18.22	-20.85	-15.59

Source: Reviewer.

Figure 6: Placebo Subtracted Change from Baseline in MADRS Total Score (TRD3001: FAS)



Source: Reviewer.

Reviewer's Results

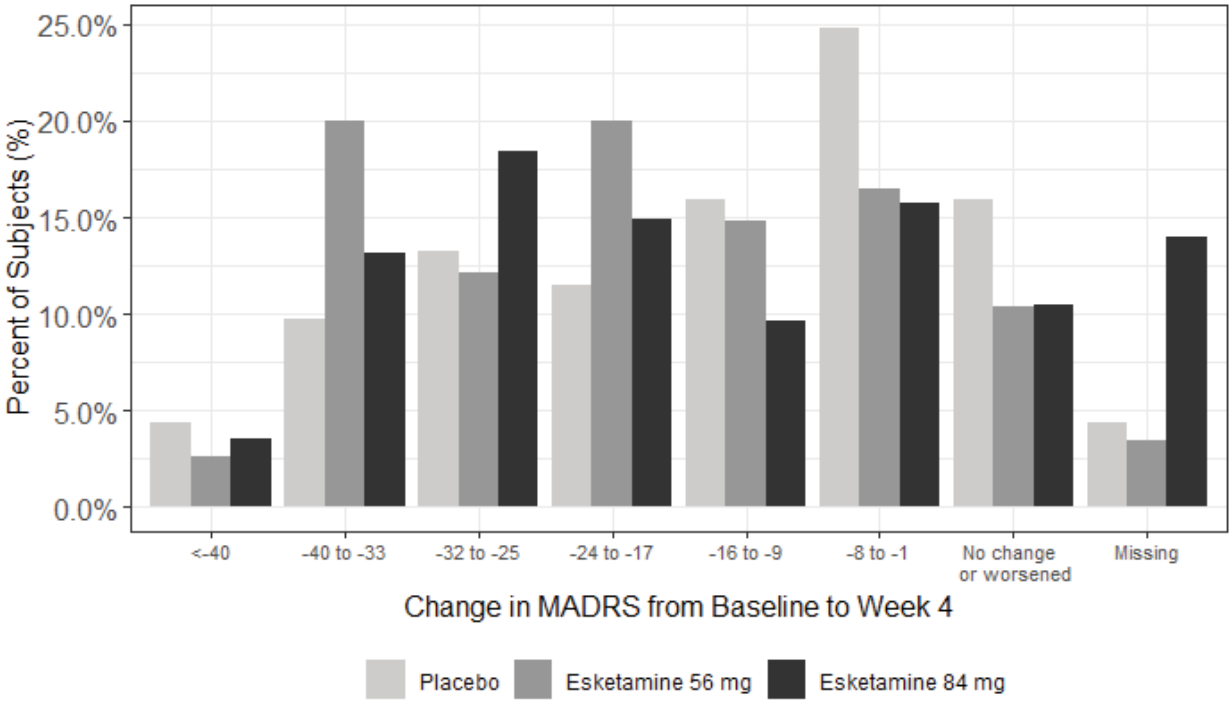
This Reviewer recreated the results for the primary analysis in TRD3001. In addition, missing data and response distributions are explored; information from TRD3001 about the dose-response is reviewed; and the time course of the treatment effect estimates and p-values are explored.

Missing Data and Response Distributions

In TRD3001, 92% of patients completed the double-blind phase. Using Figure 7, this Reviewer analyzed the usefulness of esketamine in treating TRD. The histogram includes categories for patients who showed no improvement or worsening TRD. In addition, patients who dropped out before the fourth week are included in the "Missing" category. Individual esketamine doses' usefulness is unclear from Figure 7 because no clear trends are observed (black bars – 84 mg, grey bar – 56 mg, light grey bar – placebo). However, both esketamine doses have more patients with at least a 17-point decline compared to placebo. Therefore, esketamine may be more useful than placebo.

Reference ID: 4392083

Figure 7: Percentage of Subjects with Specified Change in MADRS Total Score (TRD3001: FAS)

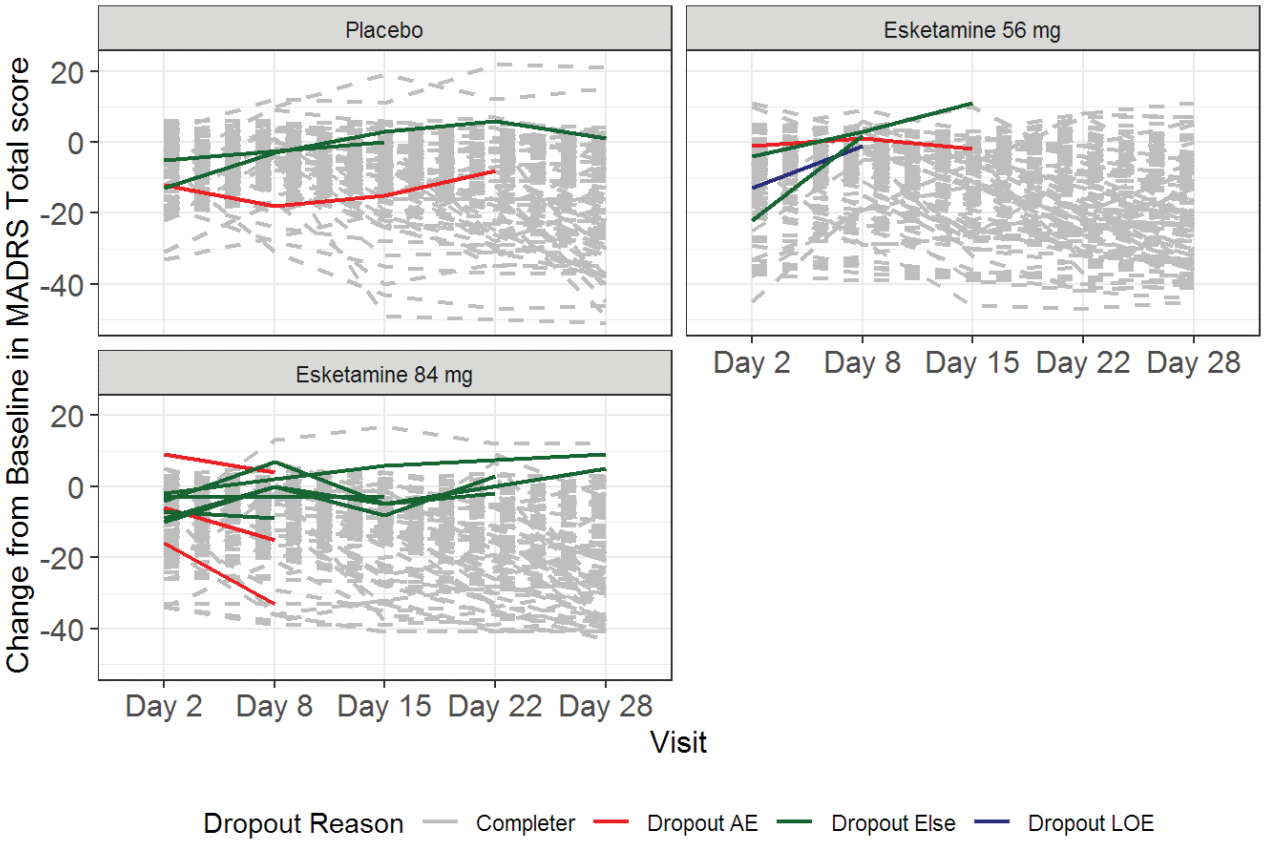


Source: Reviewer

Patients MADRS trajectories are plotted by windowed study visit in Figure 8. Dropouts due to an adverse event are red, lack of efficacy dropouts are blue, other reason dropouts are green, and completers are dashed grey. The small number of dropouts in each arm limit the potential impact of missing data. However, both dropout due to lack of efficacy and other reasons show an increasing MADRS score before dropout. This pattern may indicate informative dropout.

Reference ID: 4392083

Figure 8: Patient Level MADRS Change Score Trajectories (TRD3001: FAS).



Source: Reviewer

<u>Treatment Effect and p-Value Trajectories</u>

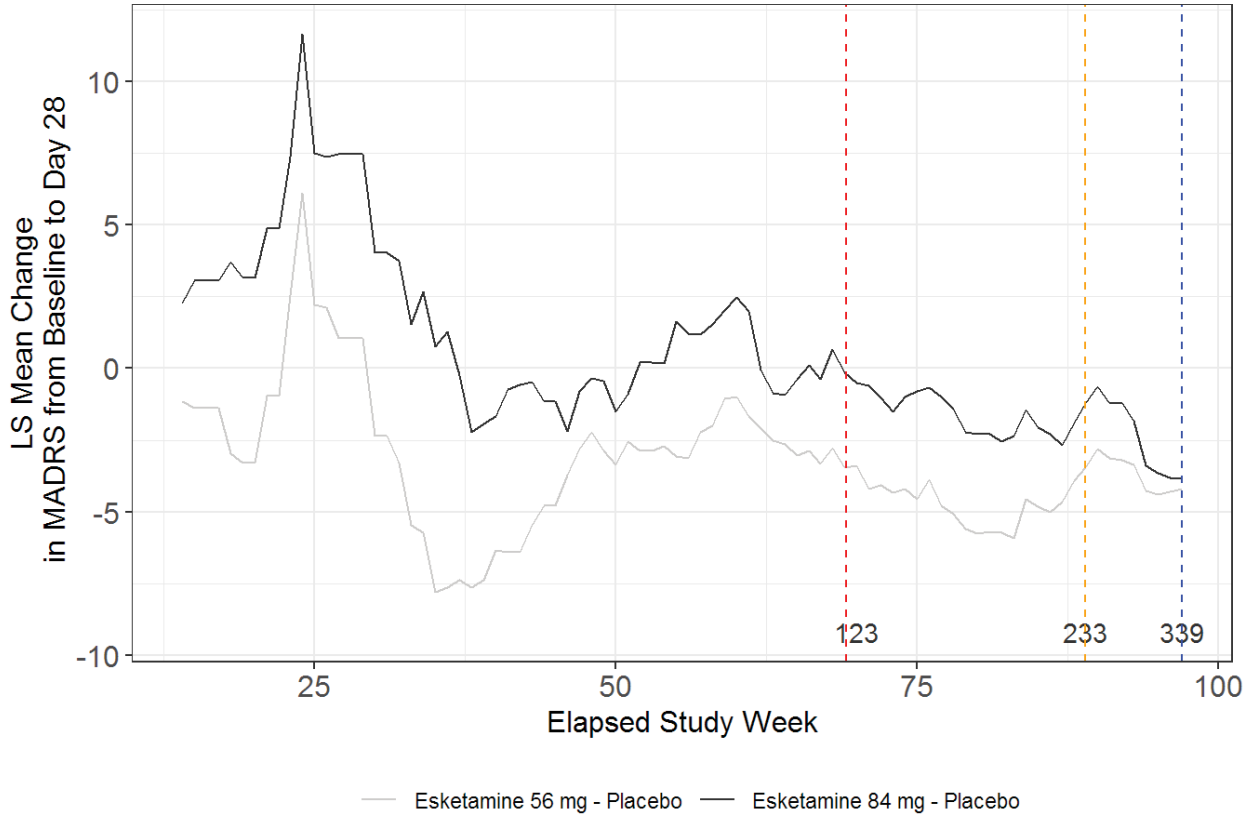
As exploration of the study conduct, this Reviewer explored the evolution of the treatment effect (difference in average MADRS score between esketamine and placebo at week 4) and p-value for the treatment effect by elapsed study week. A negative number is an improvement. The analysis started at week 14 after enough patients had finished day 28 for the MMRM to converge. In Figure 9 and Figure 10, the black line represents the 84 mg dose arm trajectory, the light grey line represents the 56 mg dose arm trajectory, the red dashed is the time of the interim analysis, the orange dashed line is time when the TRD3001 reached its minimum sample size (n = 234), the blue line indicates the final analysis, and the numbers are the sample sizes at each analysis. In Figure 9, the 84 mg dose arm always shows a worse treatment effect than the 56 mg dose arm for the entire study duration. At the interim analysis, the 84 mg dose arm shows no treatment effect. In Figure 10, the 84 mg dose arm p-value does not stabilize during the study. It trends downwards after the interim analysis potentially indicating that a larger sample size would have showed a statistically significant treatment effect. The 56 mg dose arm p-value stabilizes before the interim analysis. The downward trend in this p-value started about week 60.

<u>Reviewer's Note:</u> *These figures support the belief that the 56 mg dose arm would have had statistical support if the Applicant had reserved alpha for testing this dose arm, but it would still not have the statistical support at the time when the minimum sample size was reached (one-*

Reference ID: 4392083

sided p-value = 0.058) because of both smaller sample size and smaller observed treatment effect.

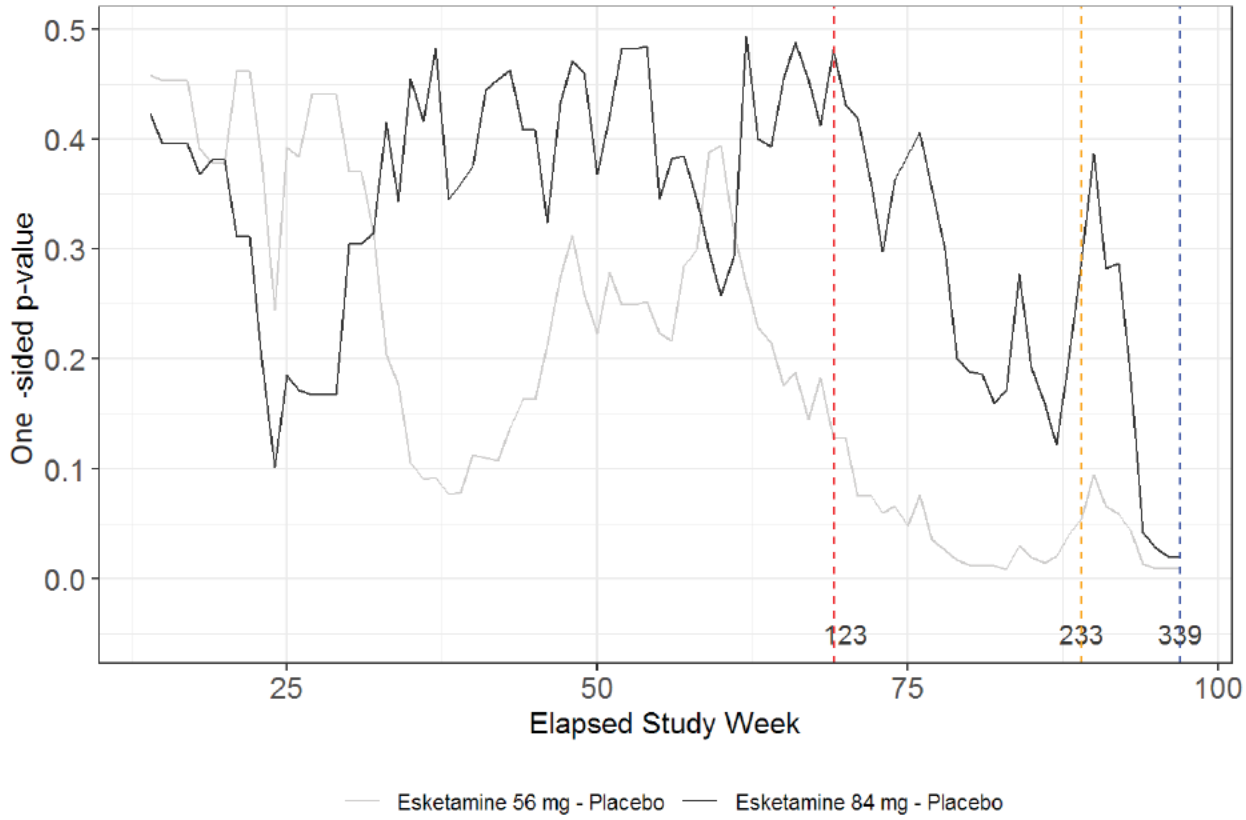
Figure 9: Esketamine LS Mean Treatment Effect over Study Elapsed Week (TRD3001: FAS)



Source: Reviewer.

Reference ID: 4392083

Figure 10: Esketamine One-sided p-value for Treatment Effect over Study Elapsed Week (TRD3001: FAS)



Source: Reviewer.

3.2.5 Study ESKETINTRD3002

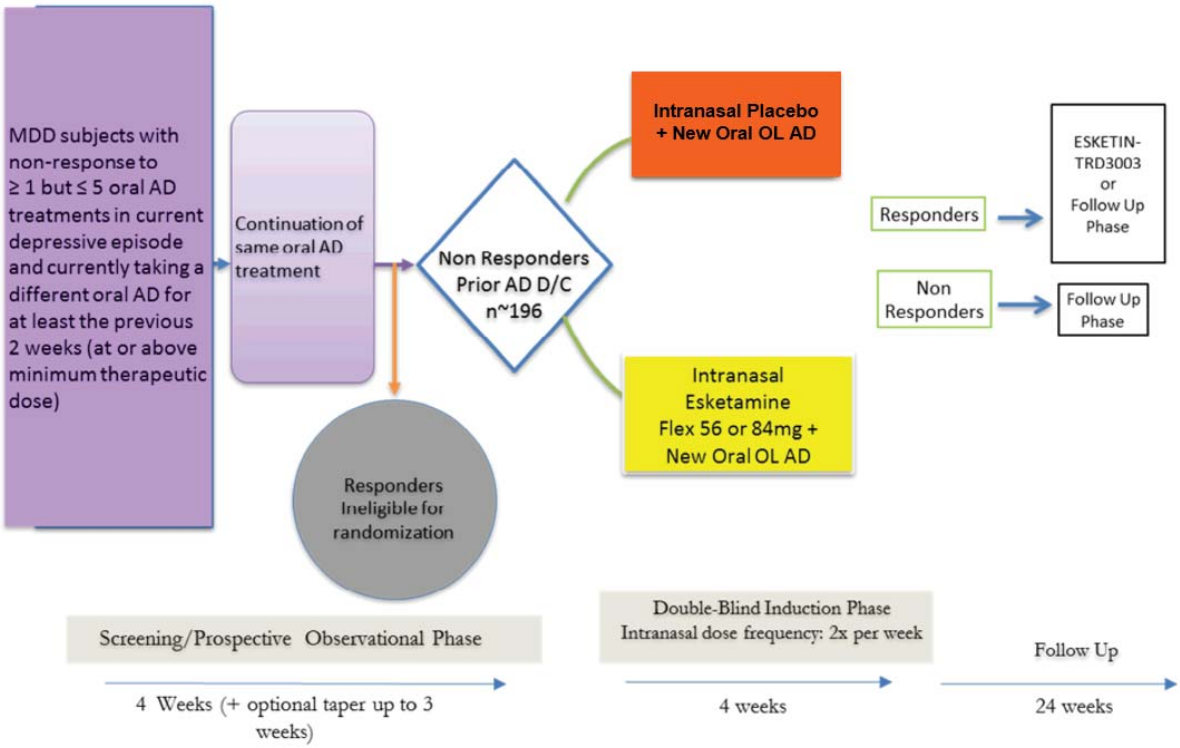
3.2.5.1 Study Design and Endpoints

TRD3002 was a flexible dose study comparing flexibly dosed esketamine 56 mg to 84 mg + OAD to placebo + OAD. All patients in the esketamine arm started on 56 mg esketamine. Doses could be increased to 84 mg esketamine on visits day 4 – day 15 if the investigator believed that a higher dose would improve patient response. Dose could only be adjusted after day 15 for tolerability reasons. The study was conducted at 39 sites worldwide with 10 sites in the United States. Subjects were randomized at a 1:1 ratio to either esketamine or placebo. The Applicant enrolled a total of 227 patients with 223 patients contributing data to the efficacy evaluation. Key secondary endpoints were:

- Onset of clinical response by Day 2
- Functioning and associated disability measured by Sheehan Disability Scale (SDS)
- Subject reported depressive symptoms measured by PHQ-9

Reference ID: 4392083

Figure 11: Study Schematic for ESKETINTRD3002



Source: Adapted from Applicant's Protocol for TRD3002.

3.2.5.2 Statistical Methodologies

<u>Applicant's Methods</u>

Study TRD3002 was designed to test flexibly dosed esketamine vs. placebo controlling for multiple comparisons using a fixed sequence procedure. The FWER was controlled over the primary and key secondary endpoints using a fixed sequence method. The Applicant designed TRD3002 to have 90% power at a sample size of 196 with a one-sided alpha level of 2.5% and 25% dropout rate. The Applicant further assumed that the treatment difference between placebo and each esketamine arm was 6.5 point change on the MADRS at week 4 with a standard deviation of 12.

The primary analysis is outlined in Section 3.2.2. The Applicant analyzed the day 2 response rate using a Cochran-Mantel-Haenszel chi-square test with strata of country and OAD class.

The Applicant also conducted a tipping point analysis (TPA). TPA examines the sensitivity of the primary MMRM analysis to violation of MAR where the dropouts have a smaller treatment effect than the completers by a penalty. The conducts three different TPA's:

Reference ID: 4392083

- [redacted] (b) (4) $to\ \Delta_1^*$ in increments of 1 (active-only adjustment analysis)
- [redacted] (b) (4) $to\ \Delta_2^*$ in increments of 1 (analysis with a control adjustment that is half of the active adjustment)
- [redacted] (b) (4) $to\ \Delta_3^*$ in increments of 1 (all arms adjustment analysis).

Source: Applicant's SAP.

Here, δ_A is the penalty for the esketamine arm, δ_A is the penalty for the placebo arm, and Δ_j^*, $j = 1,2,3$ is the value of the penalty when the results change from statistically significant to no longer statistically significant. For each penalty term, 500 datasets are imputed, analyzed with MMRM, results combined with Rubin's Rule, and p-values reported. TPA is interpreted to show the largest reduction in treatment effect before the p-value crosses a threshold (0.025 for this study).

3.2.5.3 Patient Disposition, Demographic and Baseline Characteristics

Of the 227 randomized subjects, 111 subjects were randomized to placebo, 116 subjects were randomized to flexibly dosed esketamine. The FAS consisted of 109 patients on placebo and 114 patients on esketamine. In the esketamine arm, 15% of patients dropped out compared to 10% in the placebo arm. In addition, only one placebo patient dropped out because of an adverse event compared to eight in the esketamine arm. See Table 10 for additional details. The Applicant excluded site PL10002 (nine patients) because of violations of Good Clinical Practice including an inability to link patients with administered treatment. The Applicant conducted a sensitivity analysis including these nine patients.

Reviewer's Note: *In the Applicant's sensitivity analysis including the nine patients (site PL10002), the estimated treatment difference increased in magnitude to 4.2 from 3.98 and the p-valued decreased to 0.012 from 0.0225.*

Table 10: Subject Disposition (TRD 3002)

	Placebo – N(%)	Esketamine – N(%)	Overall – N(%)
Randomized	111	116	227
Full Analysis Set	109	114	223
Completed 28-day treatment	99 (90.83)	98 (85.96)	197 (88.34)
Dropout – Adverse Event	1 (0.92)	8 (7.02)	9 (4.04)
Dropout – Lack of Efficacy	0 (0)	2 (1.75)	2 (0.90)
Dropout – Other Reasons	9 (8.26)	6 (5.26)	15 (6.73)

Source: Reviewer.

In the FAS, the patients were majority female (62%). In addition, the FAS was predominantly white (93%). Patients were an average age of 45.7 years. Patients were overweight with a mean BMI of 28 kg/m^2. Approximately 40% of patients were from the United States. Demographic variables were well balanced across all treatment arms. Patients in the esketamine arm were more likely to have only failed one previous antidepressant therapy compared to placebo (18% vs. 10%). See Table 11 for additional demographic characteristics.

Reference ID: 4392083

Table 11: Demographic Characteristics (TRD3002: FAS)

Characteristic	Statistics	Esketamine N = 114	Placebo N = 109	Overall N = 223
Age (years)	Mean	44.95	46.40	45.66
	SD	12.58	11.14	11.89
	Median	45.00	47.00	47.00
	Min, Max	19, 64	20, 64	19, 64
Sex, n (%)				
Male		39 (34)	46 (42)	85 (38)
Female		75 (66)	63 (58)	138 (62)
Race, n (%)				
American Indian or Alaska Native		0	0	0 (0)
Asian		1 (1)	1 (1)	2 (1)
Black or African American		6 (5)	5 (5)	11 (5)
Multiple		1 (1)	1 (1)	2 (1)
Not Reported		0	0	0
Other		0	0	0
White		106 (93)	102 (94)	208 (93)
Ethnicity, n (%)				
Hispanic or Latino		5 (4)	7 (6)	12 (5)
Not Hispanic or Latino		109 (96)	102 (94)	211 (94)
Weight (kg)	Mean	79.30	82.67	80.95
	SD	20.14	19.47	19.84
	Median	73.10	84.90	79.70
	Min, Max	48.90, 162.80	45.30, 147.00	45.30, 162.80
Height (cm)	Mean	169.23	169.81	169.51
	SD	10.18	9.95	10.05
	Median	168.15	167.00	168.00
	Min, Max	148.50, 193.00	151.00, 194.00	148.50, 194.00
BMI (kg/m²)	Mean	27.53	28.60	28.05
	SD	5.84	6.24	6.05
	Median	26.90	28.20	27.30
	Min, Max	16.10, 56.30	18.10,52.70	16.10, 56.30
Country, n(%)				
Czech Republic		30 (26)	28 (26)	58 (26)
Germany		10 (9)	10 (9)	20 (9)
Spain		9 (8)	9 (8)	18 (8)

Reference ID: 4392083

Poland	20 (18)	18 (17)	38 (17)
United States	45 (39)	44 (40)	89 (40)
Novel Oral Antidepressant, n(%)			
Duloxetine	60 (53)	61 (56)	121 (54)
Escitalopram	21 (18)	17 (16)	38 (17)
Sertraline	16 (14)	16 (15)	32 (14)
Venlafaxine XR	17 (15)	15 (14)	32 (14)
Failed Oral Antidepressant Therapies, n (%)			
1	10 (9)	18 (17)	28 (13)
2	69 (61)	54 (50)	123 (55)
3 or more	35 (31)	37 (34)	72 (32)

Source: Reviewer.

At baseline, subjects had an average, baseline MADRS total score of 37.27 in the placebo arm, 37.04 in the esketamine arm.

Table 12: Baseline MADRS Total Score (TRD3002: FAS)

Scale	Statistics	Placebo N = 109	Esketamine N = 114
MADRS	Mean	37.27	37.04
	SD	5.66	5.69
	Median	37	37
	Min, Max	21, 52	22, 48

Source: Reviewer.

3.2.5.4 Results and Conclusions

Applicant's Results

The primary efficacy endpoint of change from baseline to day 28 total MADRS showed a 3.98 (one-sided p = 0.0225) point greater decline in the esketamine arm compared to placebo, see Table 13. The esketamine arm declined 21 points from an average baseline score of 37 points. See Table 13 for the placebo arm values. No statistically significant difference in response rates at day 2 was observed, see Table 14. Formal hypothesis testing stopped after this endpoint. P-values were compared to an alpha level of 0.05. Throughout this Section, negative change indicates improvement.

Reference ID: 4392083

Table 13: MADRS Total Score at 28 Days (TRD3002: FAS)

	Placebo		Esketamine			
Dose	N	LS Mean (SE)	N	LS Mean (SE)	LS Mean Difference (95% CI)	Two-sided p-value
56 mg – 84 mg	109	-16.84 (1.25)	114	-20.81 (1.24)	-3.98 (-7.46, -0.50)	0.0225

Source: Reviewer.

Table 14: Applicant's Analysis of Response Rate at Day 2 (TRD3002: FAS)

	Intranasal Esk + Oral AD (N=114)	Oral AD + Intranasal Placebo (N=109)
Onset of clinical response, n (%) (a)		
N	114	109
Yes	9 (7.9%)	5 (4.6%)
No	105 (92.1%)	104 (95.4%)
Generalized Cochran-Mantel-Haenszel test (b)		
1-sided p-value (esk+AD vs. AD+placebo)	0.161 (c)	
Odds ratio (95% CI) (d)	1.79(0.57,5.67)	

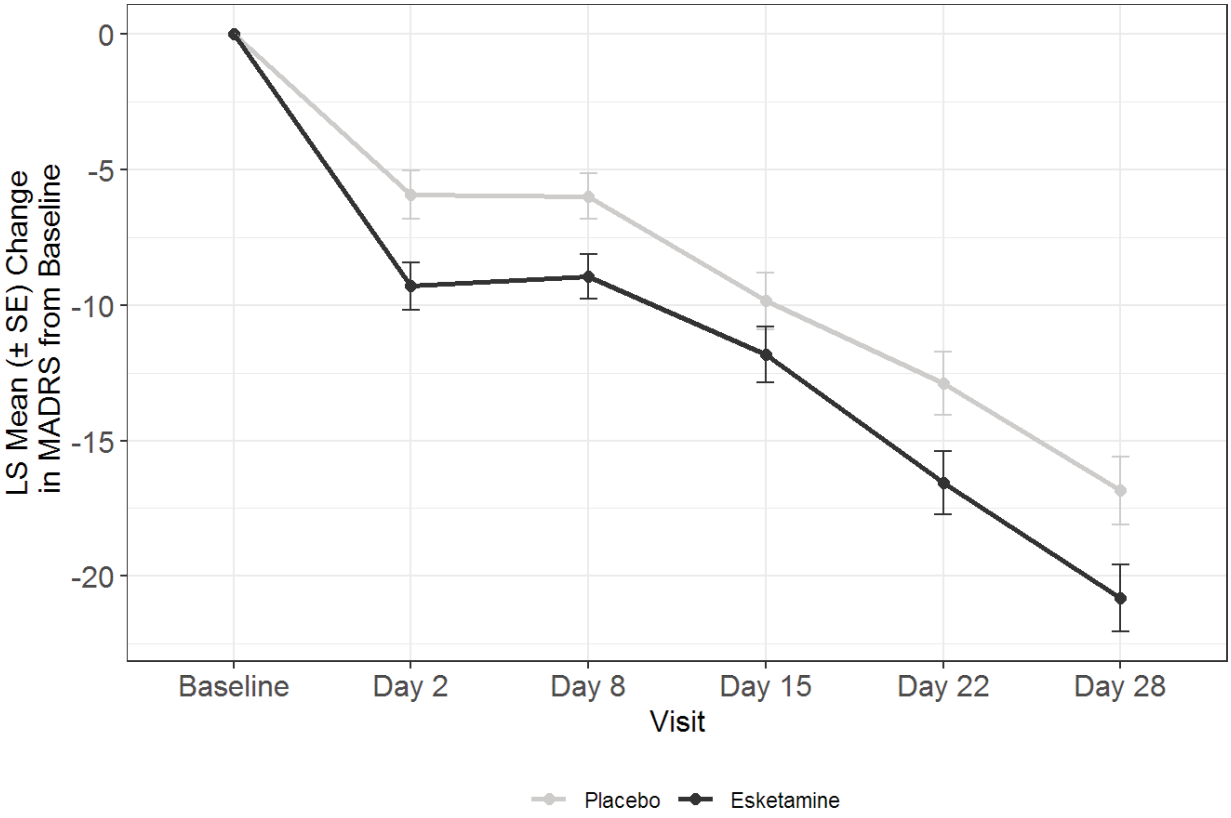
(a) Onset of clinical response is defined as at least 50% improvement from baseline in MADRS total score with onset by Day 2 that is maintained to Day 28. Subjects are allowed one excursion (non-response) on Days 8, 15 or 22, provided the score is at least 25% improvement. Subjects with missed assessments or discontinued early are not considered to have onset of clinical response.
(b) Generalized Cochran-Mantel-Haenszel (CMH) test for mean score difference between treatments adjusting for country and class of oral antidepressant (SNRI or SSRI).
(c) The analysis can be considered statistically significant at the 1-sided 0.025 level only if the change in MADRS total score analysis is also significant. If not statistically significant, the onset of clinical response cannot be formally evaluated and the p-value should not be referenced.
(d) Odds of achieving onset of clinical response on intranasal esk + oral AD divided by the odds of achieving onset of clinical response on oral AD + intranasal placebo.

Source: Applicant's TRD3002 Clinical Study Report.

Over the 28-day study period, the average MADRS total score declined in both the placebo and esketamine arms, see Figure 12. The esketamine arm (black line) showed a greater decline than placebo in the first 24 hours after administration of the first dose of esketamine. The initial difference remained throughout the 28 days of double blind therapy, see Figure 13.

Reference ID: 4392083

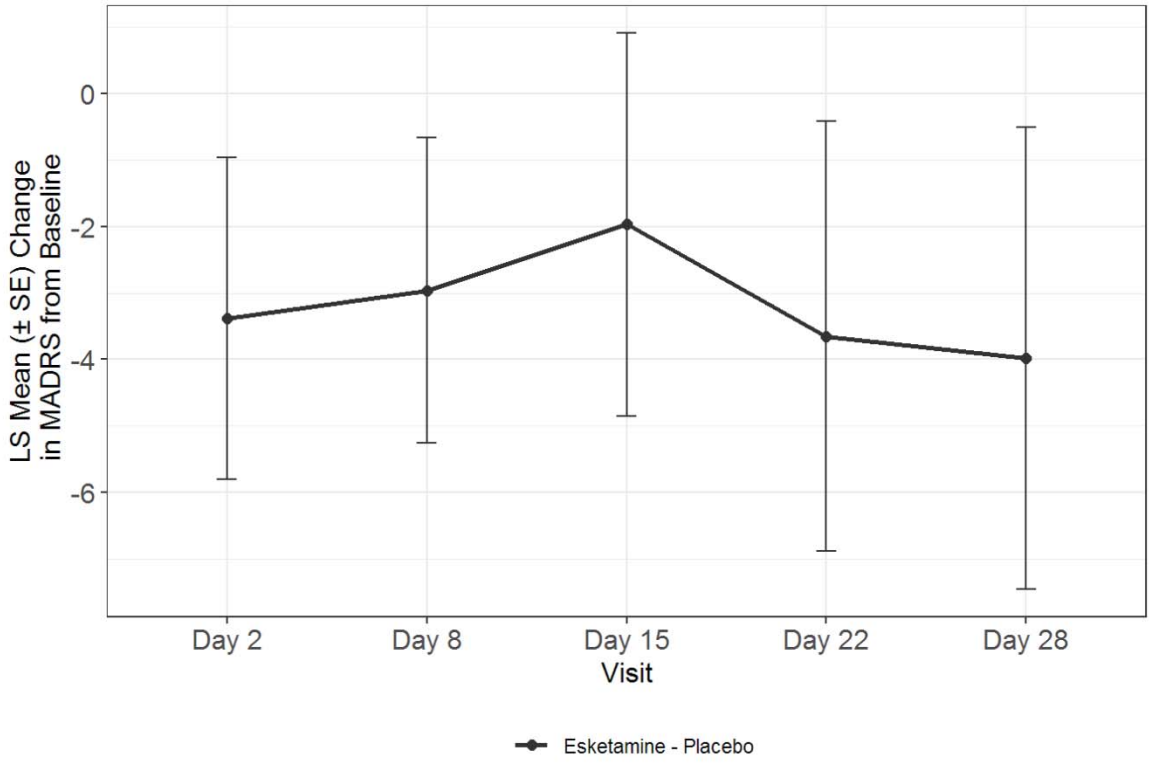
Figure 12: MADRS Total Score - LS Mean (± SE) Change from Baseline over Time - Mixed Model for Repeated Measures (TRD3002: FAS)



Source: Reviewer

Reference ID: 4392083

Figure 13: MADRS Total Score - LS Mean (± SE) Treatment Effect over Time - Mixed Model for Repeated Measures (TRD3002: FAS)



Source: Reviewer.

Table 15: Change from Baseline in MADRS Total Score (TRD3002: FAS)

Visit	*Treatment Arm*	*Change from Baseline*	*SE*	*Lower 95% CI*	*Upper 95% CI*
Day 2	Placebo	-5.92	0.89	-7.67	-4.17
	Esketamine	-9.30	0.87	-11.00	-7.60
Day 8	Placebo	-5.98	0.84	-7.64	-4.32
	Esketamine	-8.94	0.83	-10.57	-7.31
Day 15	Placebo	-9.84	1.05	-11.91	-7.77
	Esketamine	-11.81	1.03	-13.83	-9.78
Day 22	Placebo	-12.90	1.17	-15.19	-10.60
	Esketamine	-16.55	1.16	-18.82	-14.27
Day 28	Placebo	-16.84	1.25	-19.30	-14.37
	Esketamine	-20.81	1.24	-23.26	-18.37

Source: Reviewer.

Table 16: Weekly Treatment Effect in MADRS Total Score (TRD3002: FAS)

Visit	*Treatment Arm*	*Treatment Effect*	*SE*	*p-value*
Day 2	Esketamine - Placebo	-3.38	1.21	0.0052
Day 8	Esketamine - Placebo	-2.96	1.15	0.0099
Day 15	Esketamine - Placebo	-1.97	1.44	0.1729

Reference ID: 4392083

Day 22	Esketamine - Placebo	-3.65	1.62	0.0241
Day 28	Esketamine - Placebo	-3.98	1.74	0.0225

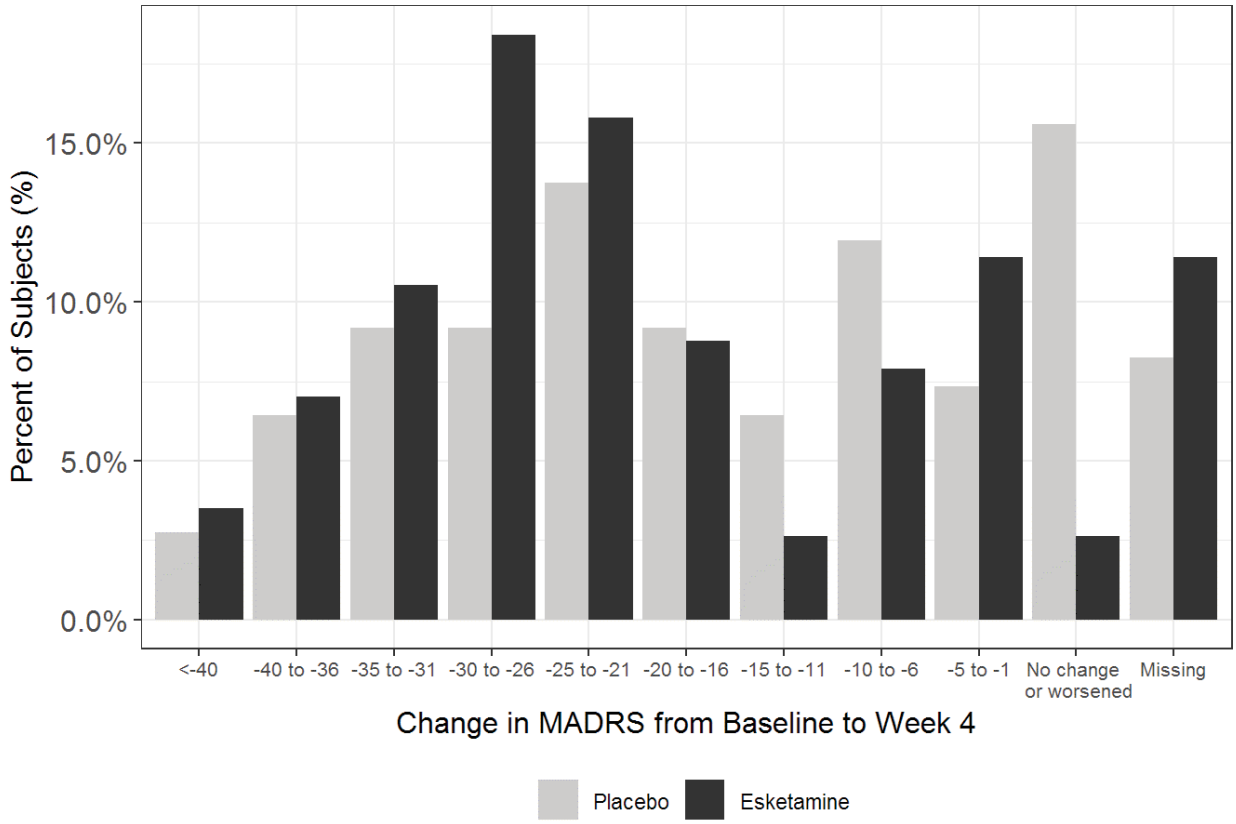
Source: Reviewer.

The Applicant's TPA only lost statistical significance in the case when esketamine dropouts did 9 points worse on the MADRS compared to completers and placebo dropouts did 0 points worse than placebo completers.

Reviewer's Results

This Reviewer recreated the results for the primary endpoint in TRD3002. Results for the key secondary endpoints have not yet been recreated. Figure 14 summarizes the response distribution of TRD patients to esketamine treatment, see description of this plot under TRD3001. In Figure 14, the esketamine arm (black bars) has a greater percentage of patients with at least a 20 point decline in MADRS scores. In addition, the placebo arm as greater percentage (>15%) of patients showing no change or worsening compared to ~2.5% of patients in the esketamine arm. Combined with comparable dropout rates across arms implies that esketamine is useful in treating TRD.

Figure 14: Percentage of Subjects with Specified Change in MADRS Total Score (TRD3002: FAS)

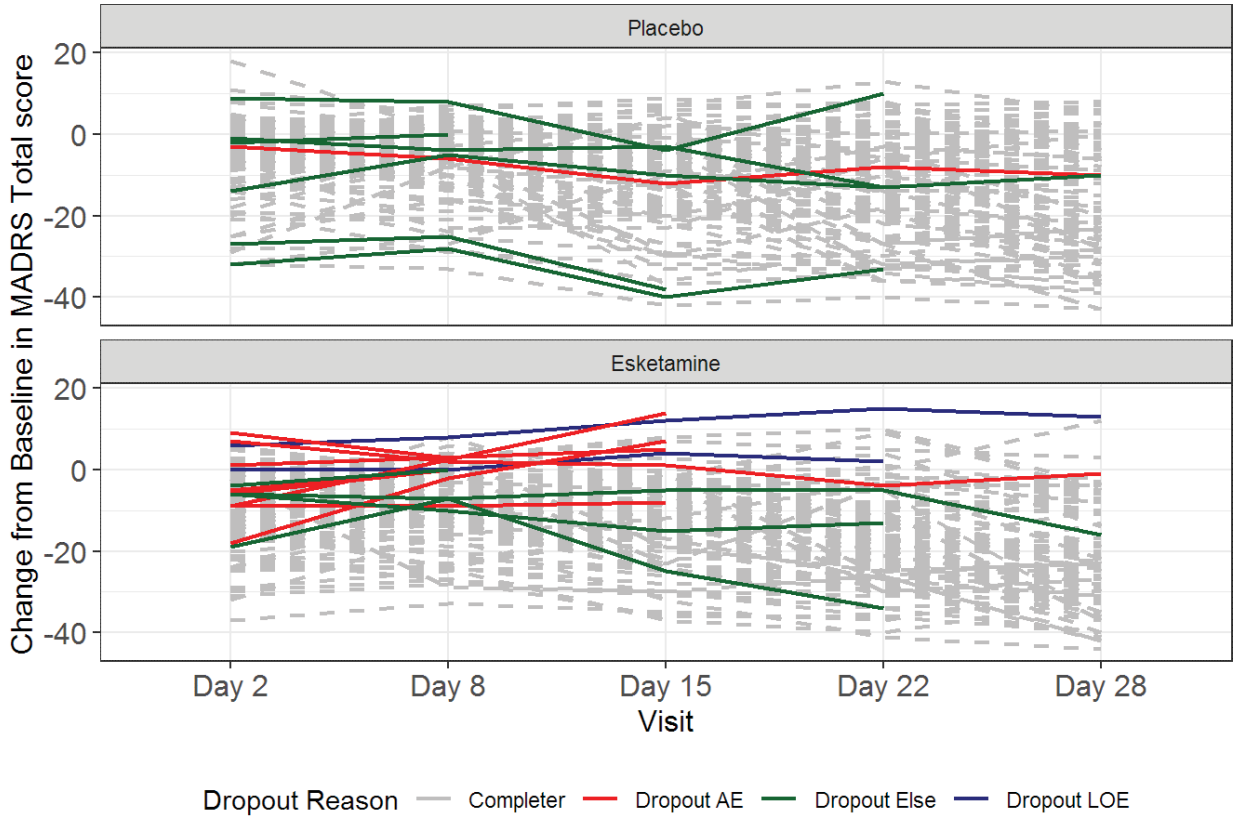


Source: Reviewer

Reference ID: 4392083

In Figure 15, patient specific MADRS total score trajectories are plotted against study visit. Dropouts due to an adverse event are red, lack of efficacy dropouts are blue, other reason dropouts are green, and completers are dashed grey. In both treatment arms, dropouts do not appear to differ systematically from completers. There is no evidence against assuming MAR missing data mechanism.

Figure 15: Patient Level MADRS Change Score Trajectories (TRD3002: FAS)



Source: Reviewer.

In TRD3002, patients were more likely end the study on 84 mg esketamine every three days (68%) compared to 32% of patients on 56 mg esketamine, see Table 17. Sample patient specific, esketamine dose trajectories are presented in Figure 16. In six of sample trajectories, the esketamine dose was titrated on day 4. One patient's dose was reduced on day 8. Three patients remained on 56 mg through the double-blind period. Any conclusions about the dose-response using information from TRD3002 must consider that dose adjustments depend on the patient's treatment response and tolerability. Often, a patient receives a higher dose of esketamine because they are not improving or they can tolerate the higher dose.

Table 17: Number of Patients per Dose Level (TRD3002: FAS)

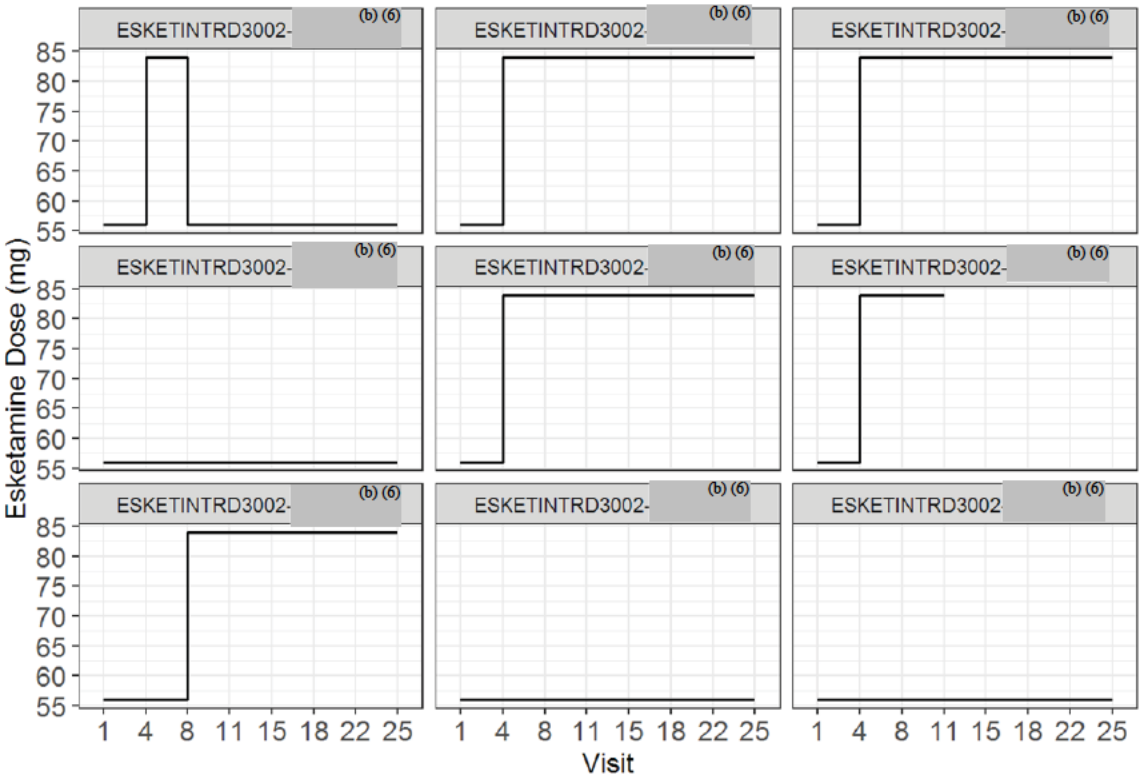
	1	**4**	**8**	**11**	**15**	**18**	**22**	**25**
Day	N (%)							
56 mg esketamine	114 (100)	58 (54)	40 (37)	30 (28)	34 (32)	34 (33)	33 (32)	32 (32)

Reference ID: 4392083

84 mg esketamine	0	49 (46)	68 (63)	77 (72)	71 (68)	68 (67)	69 (68)	67 (68)
Placebo	109 (100)	106 (100)	105 (100)	103 (100)	101 (100)	100 (100)	101 (100)	98 (100)

Source: Reviewer.

Figure 16: Sample Esketamine Dose Trajectories (TRD3002: FAS)



Source: Reviewer.

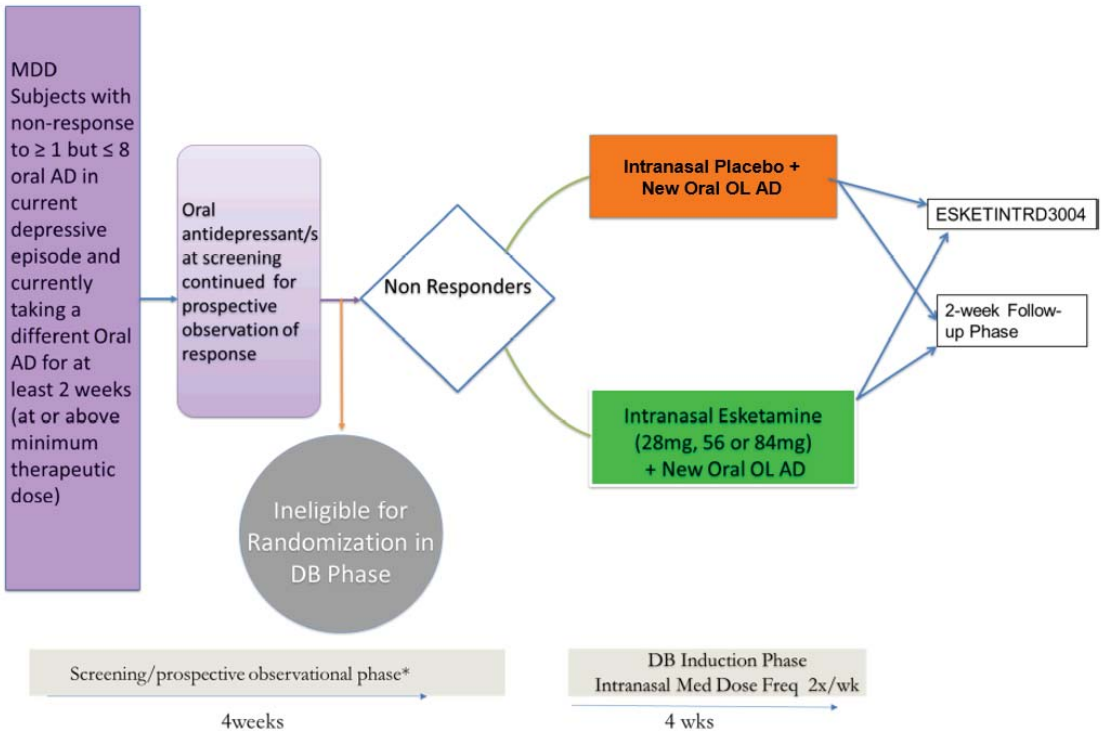
3.2.6 Study ESKETINTRD3005

3.2.6.1 Study Design and Endpoints

TRD3005 was flexible dose study comparing flexibly dosed esketamine (28 mg, 56 mg, or 84 mg) + OAD to placebo + OAD. All patients in the esketamine arm started on 28 mg esketamine. Doses could be increased to 56 mg or 84 mg esketamine on visits day 4 visit if the investigator believed that a higher dose would improve patient response. Doses may be reduced by 28 mg per study visit at visits after day 15 if needed for tolerability. The Applicant planned to originally enroll 100 patients with a sample size re-estimation after 50 patients were enrolled. The maximum potential sample size after re-estimation was 148 patients. TRD3005 only enrolled patients 65 years old and older.

Reference ID: 4392083

Figure 17: Study Schematic for ESKETINTRD3005



Source: Adapted from Applicant's CSR.

3.2.6.2 Statistical Methodologies

Applicant's Methods

The Applicant analyzed the primary endpoint using the same methodology as TRD3001. See Section 3.2.4.2 for details. The Applicant only planned to control type I error on the primary endpoint. The sample size re-estimation criteria are in Table 18.

Table 18: Sample Size Re-Estimation Rules (TRD3005)

Conditional Power (%)	p-value (1-sided)	Decision	Re-estimated total sample size
<1	>0.41	Stop study	
30<CP<50	--	Continue study	148
50≤CP<80	--	Continue study	124
1≤CP≤30 or CP≥80	--	Continue study	100

Source: Applicant's SAP for TRD3005.

3.2.6.3 Interim Analysis Results

The preplanned interim analysis was conducted after 51 patients were randomized on 15 February 2017. The DMC recommended continuing the study with a final sample size of 100 patients. The Applicant decided to enroll 137 patients instead of the recommended 100 patients because the Applicant again decided to allow all patients in screening phase to enter the double-blind period if they qualify.

Reference ID: 4392083

3.2.6.4 Patient Disposition, Demographic and Baseline Characteristics

In TRD3005, 66 patients were randomized to placebo and 72 patients were randomized to esketamine. The FAS consisted of 137 patients. Across both arms, 89% of patients completed the 4 week double-blind phase. Additional details of patient disposition are presented in Table 19.

Table 19: Subject Disposition (TRD3005)

	Placebo – N(%)	Esketamine – N(%)	Overall – N(%)
Randomized	66	72	138
Full Analysis Set	65	72	137
Completed 28-day treatment	60 (92.31)	62 (86.11)	122 (89.05)
Dropout – Adverse Event	2 (3.08)	4 (5.56)	6 (4.38)
Dropout – Lack of Efficacy	1 (1.54)	3 (4.17)	4 (2.92)
Dropout – Other Reasons	2 (3.08)	3 (4.17)	5 (3.65)

Source: Reviewer.

In the FAS, most the patients were female (38% male vs. 62% female). Patients had a median age of 69 years and ranged from 65 to 86 years. The patient population was majority white (95%). See Table 20 for additional demographic characteristics.

Table 20: Demographic Characteristics (TRD3005: FAS)

Characteristic	Statistics	Esketamine N = 72	Placebo N = 65	Overall N = 137
Age (years)	Mean	70.62	69.35	70.02
	SD	4.79	4.15	4.52
	Median	70.00	68.00	69.00
	Min, Max	65, 86	65, 82	65, 86
Sex, n (%)				
Male		27 (38)	25 (38)	52 (38)
Female		45 (62)	40 (62)	85 (62)
Race, n (%)				
Black or African American		0	0	0
Multiple		4 (6)	0	4 (3)
Not Reported		1 (1)	1 (2)	2 (1)
Other		1 (1)	0	1 (1)
White		66 (92)	64 (98)	130 (95)
Ethnicity, n (%)				
Hispanic or Latino		10 (14)	5 (8)	15 (11)

Reference ID: 4392083

Not Hispanic or Latino		62 (86)	60 (92)	122 (89)
Weight (kg)	Mean	78.44	81.70	79.99
	SD	16.67	18.95	17.79
	Median	76.25	77.60	77.50
	Min, Max	44.00, 113.50	43.00, 134.80	43.00, 134.80
Height (cm)	Mean	165.46	166.79	166.09
	SD	10.95	10.35	10.65
	Median	162.80	165.00	164.00
	Min, Max	139.70, 191.00	149.90, 193.00	139.70, 193.00
BMI (kg/m²)	Mean	28.57	29.32	28.92
	SD	5.17	6.14	5.64
	Median	27.60	28.30	27.80
	Min, Max	20.30, 45.30	16.30, 45.20	16.30, 45.30
Country, n(%)				
Belgium		2 (3)	4 (6)	6 (4)
Bulgaria		3 (4)	0	3 (2)
Brazil		1 (1)	0	1 (1)
Spain		4 (6)	4 (6)	8 (6)
Finland		1 (1)	1 (2)	2 (1)
France		4 (6)	3 (5)	7 (5)
United Kingdom		1 (1)	0	1 (1)
Italy		6 (8)	3 (5)	9 (7)
Lithuania		2 (3)	0	2 (3)
Poland		4 (5)	3 (5)	7 (5)
Sweden		8 (11)	6 (9)	14 (10)
United States		34 (47)	36 (55)	70 (51)
South Africa		2 (3)	5 (8)	7 (5)
Novel Oral Antidepressant, n(%)				
Duloxetine		25 (35)	23 (35)	48 (35)
Escitalopram		25 (35)	25 (38)	50 (36)
Sertraline		15 (21)	10 (15)	25 (18)
Venlafaxine XR		7 (10)	7 (11)	14 (10)

Source: Reviewer.

At baseline, patients in the esketamine arm had a MADRS score of 34.75 and the placebo patients had a MADRS score 35.53, see Table 21.

Table 21: Baseline MADRS Total Score (TRD3005: FAS)

Scale	Statistics	Placebo (N = 65)	Esketamine (N = 72)
MADRS	Mean	35.53	34.75
	SD	6.44	5.91
	Median	35	36

Reference ID: 4392083

Min, Max	19, 51	23, 50

Source: Reviewer.

3.2.6.5 Results and Conclusions

Applicant's Results

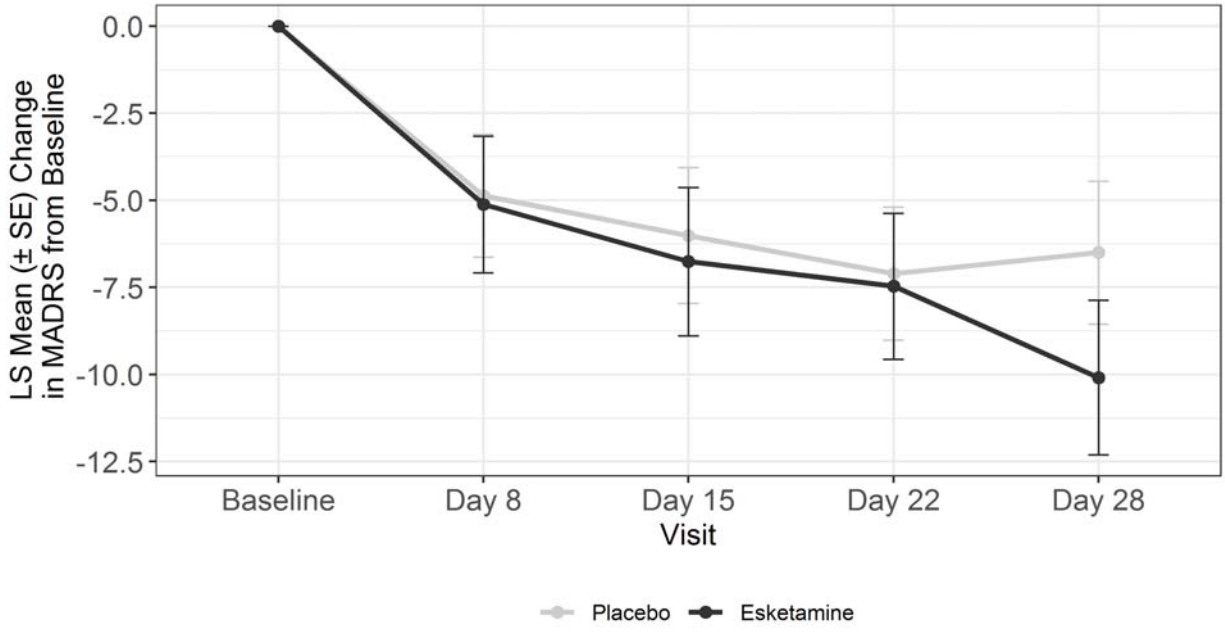
Flexibly dosed esketamine had a -3.61 (95% CI: -7.25, 0.02, p = 0.026) greater change from baseline to week 4 than placebo, see Table 22. P-values were compared to an alpha level of 0.025. The esketamine and placebo MADRS trajectories only separated at day 28, see Table 23 and Figure 18. At day 22, both esketamine and placebo arms declined by about 7 – 7.5 points, see Table 24. The placebo arm did not improve between day 22 (-7.11 points) and day 28 (-6.50 points).

Table 22: MADRS Total Score at 28 Days (TRD3005: FAS)

	Placebo		Esketamine			
Dose	N	LS Mean (95% CI)	N	LS Mean (95% CI)	LS Mean Difference (95% CI)	One-sided p-value
28 mg – 84 mg	65	-6.50 (-9.43, -3.57)	72	-10.09 (-13.05, -7.13)	-3.61 (-7.25, 0.02)	0.026

Source: Reviewer.

Figure 18: MADRS Total Score - LS Mean (± SE) Change from Baseline over Time - Mixed Model for Repeated Measures (TRD3005: FAS)



Source: Reviewer

Table 23: Weekly Treatment Effect in MADRS Total Score (TRD3005: FAS)

Treatment Arm	Visit	Treatment Effect	Lower 95% CI	Upper 95% CI	Combined one-sided p-value

Reference ID: 4392083

Esketamine - Placebo	Day 8	-0.42	-2.96	2.11	0.0596
Esketamine - Placebo	Day 15	-0.86	-3.77	2.05	0.2266
Esketamine - Placebo	Day 22	-0.34	-3.53	2.85	0.2287
Esketamine - Placebo	Day 28	-3.61	-7.25	0.02	0.0260

Source: Reviewer.

Table 24: Change from Baseline in MADRS Total Score (TRD3005: FAS)

Treatment Arm	Visit	Change from Baseline	Lower 95% CI	Upper 95% CI
Placebo	Baseline	0.00	NA	NA
Esketamine	Baseline	0.00	NA	NA
Placebo	Day 8	-4.87	-7.16	-2.59
Esketamine	Day 8	-5.12	-7.47	-2.78
Placebo	Day 15	-6.01	-8.52	-3.51
Esketamine	Day 15	-6.76	-9.29	-4.23
Placebo	Day 22	-7.11	-9.77	-4.45
Esketamine	Day 22	-7.47	-10.17	-4.78
Placebo	Day 28	-6.50	-9.43	-3.57
Esketamine	Day 28	-10.09	-13.05	-7.13

Source: Reviewer.

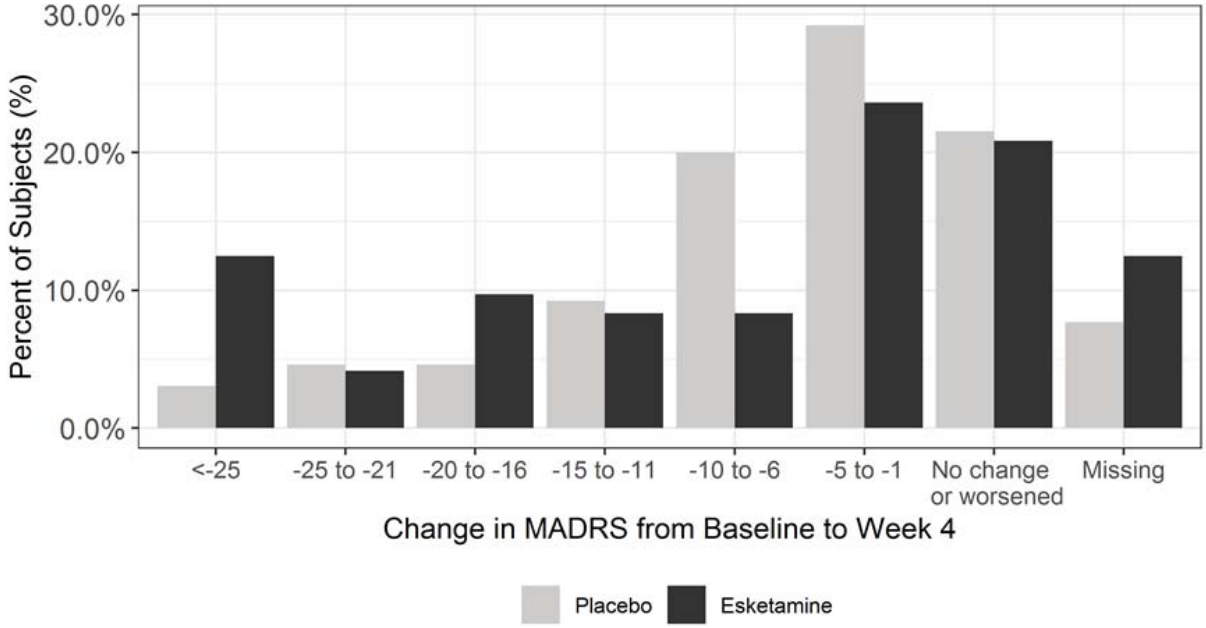
Reviewer's Results

This Reviewer recreated the results for the primary endpoint in TRD3005. Figure 19 displays the percentage of subjects with a specified change in MADRS total score at day 28, detailed description under TRD3005. The response histogram raises questions about the usefulness of esketamine in treating TRD in elderly patients. Both placebo (light grey) and esketamine (black) have a similar proportion of patients with less than a 5 point improvement.

In Figure 20, patient specific MADRS trajectories are plotted against study visit. One patient in the esketamine arm missed the pre-planned second study visit because of an increase in blood pressure to 170/110 mm Hg. This patient later dropped out on day 22. In addition, no consistent patterns are seen in the adverse event trajectories before dropout.

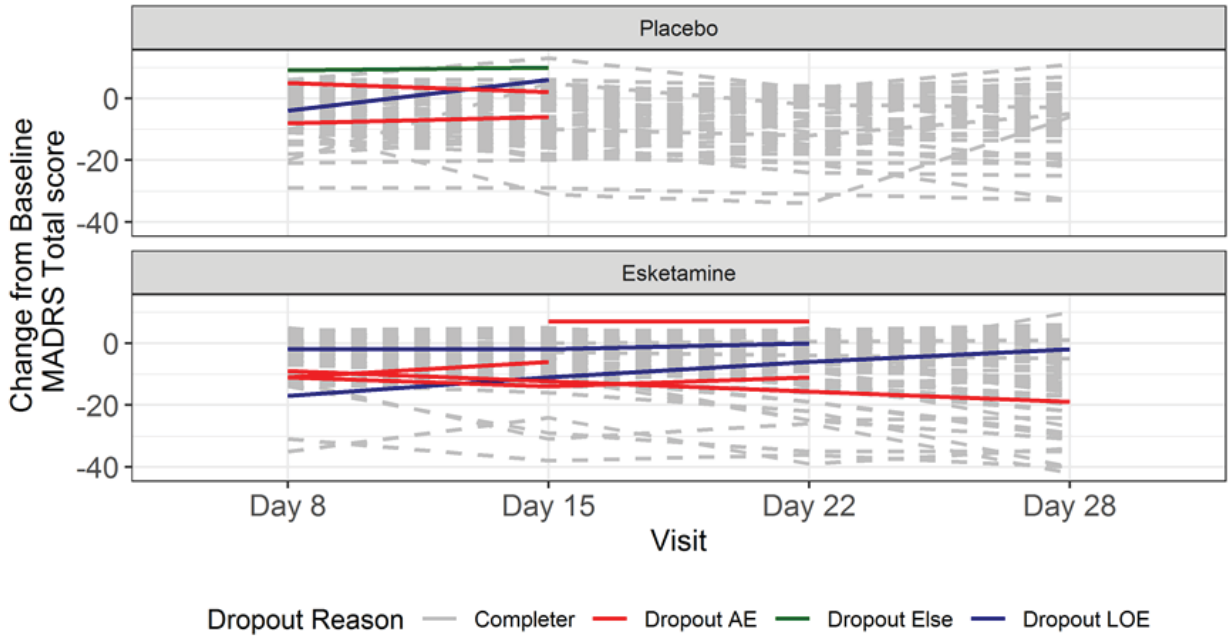
Reference ID: 4392083

Figure 19: Percentage of Subjects with Specified Magnitude of Change in MADRS Total Score (TRD3005: FAS)



Source: Reviewer

Figure 20: Patient Level MADRS Change Score Trajectories (TRD3005: FAS)



Source: Reviewer.

Reference ID: 4392083

3.2.7 Study ESKETINTRD3003

3.2.7.1 Study Design and Endpoints

TRD3003 was a randomized withdrawal study designed to demonstrate that continued use of esketamine (56 mg and 84 mg – administered once a week or once every two weeks) delays relapse of depression. Patients entered the study from either TRD3001, TRD3002, or directly into TRD3003. If patients entered from TRD3001 and TRD3002, the Applicant claims that these patients remained blinded. If patients directly entered TRD3003, they received open-label esketamine. All patients entered a twelve-week optimization period. If patients were in stable remission at the end of twelve weeks, they were randomized to switch to placebo or remain on esketamine. A second set of patients in stable response was also randomized. See Figure 21 for a detailed study schematic.

A patient reached stable remission if:

- Before protocol amendment 4: MADRS ≤ 12 for the final 4 weeks of optimization phase.

- After protocol amendment 4: allows one missed week on weeks MADRS total score ≤12 for at least 3 of the last 4 weeks of the optimization phase, but one excursion of a MADRS total score >12 or one missing MADRS assessment is permitted at optimization week 9 or 10 only. The MADRS total score at weeks 11 and 12 must be ≤12.

A patient reaches stable response if:

- Before protocol amendment 4: ≥ 50% reduction in the MADRS total score from baseline (Day 1 of induction phase; pre-randomization/prior to the first intranasal dose) in each of the last 2 weeks of the optimization phase, but does not meet criteria for stable remission with at least 1 MADRS total score of >12 in these 2 weeks.

- Before protocol amendment 4: ≥ 50% reduction in the MADRS total score from baseline (Day 1 of induction phase; pre-randomization/prior to the first intranasal dose) in each of the last 2 weeks of the optimization phase but does not meet criteria for stable remission.

Patients are followed until relapse, dropout, or end of study. A patient relapses when:

- 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.

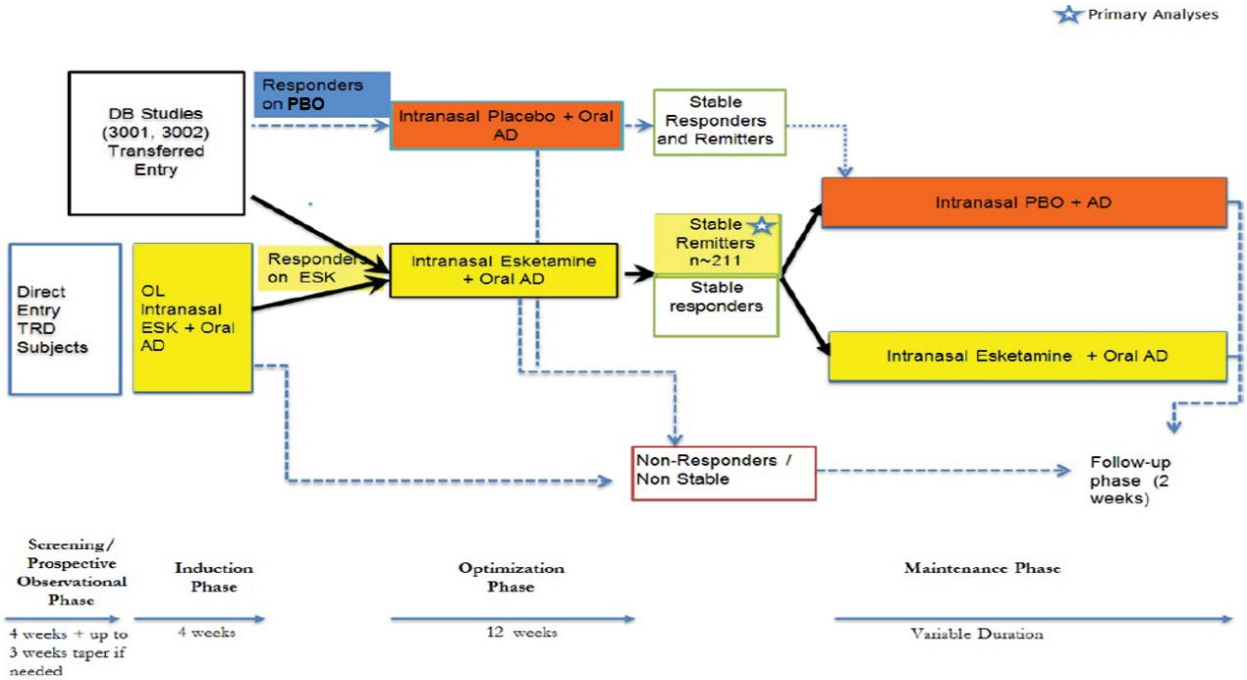
Reference ID: 4392083

- In case both relapse criteria are met, the earlier date will be defined as the date of relapse for this subject.

The Applicant planned to continue the study until at most 84 relapses are observed in the stable remitter set with an interim analysis after 30 relapses. If results of the interim analysis indicate that the study should continue, the sample size is also re-estimated with the new sample only communicated to Clinical Supply Division and no one else.

Reviewer's Note: *This Review focuses on the results from the remission set because this analysis is the primary analysis for TRD3003. Issues raised about the primary analysis for TRD3003 also apply to the secondary analysis in the stable responder set.*

Figure 21: Study Schematic for ESKETINDTRD3003



Source: Adapted from Applicant's Analysis Data Reviewer's Guide for TRD3003

3.2.7.2 Statistical Methodologies

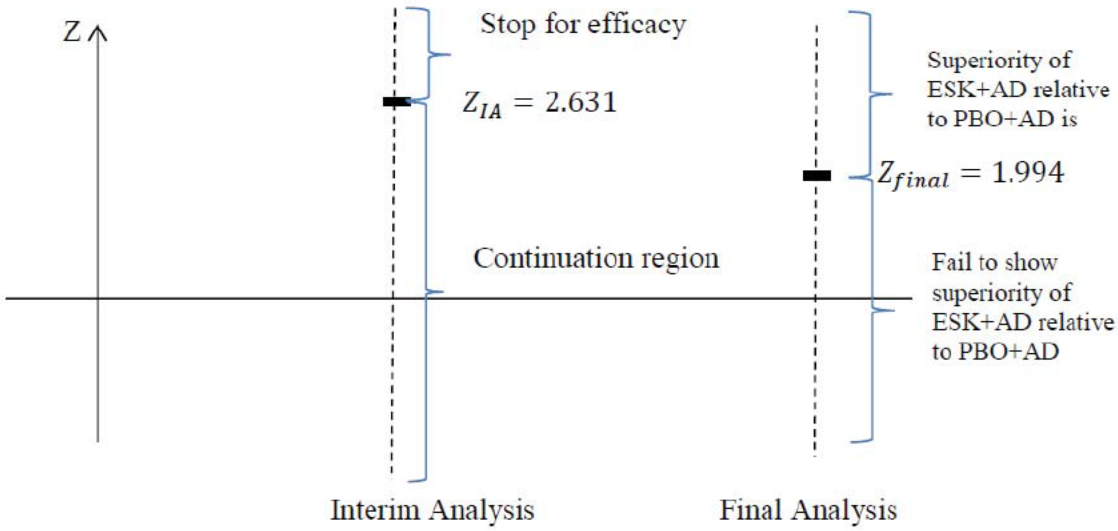
Applicant's Methods

Study TRD3003 was designed to test the treatment effect esketamine vs. placebo in delaying time to relapse. The Applicant claims that 84 events provide 90% power to detect a hazard ratio (esketamine/placebo) of 0.493 with type I error rate controlled at one-sided 2.5% level. Furthermore, for sample size calculation, the time to first relapse is assumed to follow an exponential distribution a median time to relapse of 6 months in the placebo + oral antidepressant compared to 12.17 months for the esketamine + oral antidepressant. The Applicant made additional assumptions about patient accrual period yielding at planned enrollment of 211 patients. The Applicant added an interim analysis after 30 events. They controlled type I error using a group sequential design with a single interim look (alpha spent at

Reference ID: 4392083

interim: 0.0086), see Figure 22. The sample size was re-estimated to ensure that conditional power at the final analysis is at least 90%.

Figure 22: Group Sequential Boundary (TRD3003)



Source: Applicant's Statistical Analysis Plan.

The stable remitter, full analysis set (FAS) consists of all randomized patients who are in stable remission and receive at least one dose of intranasal study drug and oral antidepressant. The Applicant conducted log-rank tests (LR) at the interim analysis and final analysis to conduct statistical inference. At the final analysis, the log-rank test is a weighted combination of the one-sided LR_1 and LR_2. LR_2 is estimated using the entire FAS, and LR_1 is estimated using the FAS consisting of all patients who received at least one dose study drugs before the interim analysis.

Figure 23: Final Log-Rank Test Statistic (TRD3003)

(b) (4)

Source: Applicant's Statistical Analysis Plan for TRD3003.

Since the study was not stopped for efficacy at the IA, the final efficacy analysis was performed

Reference ID: 4392083

at a significance level of 0.046 (2-sided).

To address concerns about changes in dissociative symptoms, the Applicant conducted a post hoc analysis excluding patients who relapsed within four weeks of randomization and had a clear change in dissociative symptoms. Dissociative symptoms are measured on the Clinician-Administered Dissociative States Scale (CADSS).

Reviewer's Note: *Changes in a patient's perception of their treatment assignment may have been influenced by esketamine acute side effects in all studies including TRD3001, TRD3002, and TRD3005. This Reviewer only includes an exploratory analysis of TRD3003 because of concerns that the optimization period may have made patients more aware of the changes in side-effects. Further research is needed for both the short term studies and randomized withdrawal studies.*

Reviewer's Methods

This Reviewer conducted additional exploratory analyses to explore the relationship between CADSS measured dissociation and relapse using a joint model for both the CADSS score trajectories over study day (longitudinal process) and time to relapse (survival process). Type I error is not controlled in these exploratory analyses. A joint model is a statistical method used when the longitudinal process and survival process are correlated within a subject. The joint model links the processes with a shared random effect. We used the JM package[2] in R developed by Dr. Dimitris Rizopoulos. While the Applicant measured CADSS scores at pre-dose, 40 minutes post-dose, and 90 minutes post-dose, only the 40 minute measurement is modeled because most of the dissociative effects are observed by this time point. Because CADSS scores have a skewed distribution with zeros, the square root transformation[3] was used to stabilize the variance and improve longitudinal model fit. The longitudinal model for the square root transformed CADSS score:

(b) (4)

where:

- (b) (4)
-
-
-

[2] Rizopoulos, D. (2010). JM: An R Package for the Joint Modelling of Longitudinal and Time-to-Event Data. *Journal of Statistical Software, 35*(9), 1 - 33. doi:http://dx.doi.org/10.18637/jss.v035.i09

[3] The log transformation has similar properties to the square transformation, but it requires addition of a small positive number when zeros are present in the data.

Reference ID: 4392083

- [redacted] (b) (4)
- [redacted]
- [redacted]

Treatment assignment was modeled using a time dependent covariate that assigned all patients to esketamine when t < 0 and assigned placebo patients to placebo o[redacted] (b) (4) [redacted] (b) (4) e fitted a linear spline model with a single knot near t_{rand}. Define [redacted] (b) (4) [redacted] (b) (4) [redacted] (b) (4). We chose a linear spline model with a single knot at t_{rand} as the simplest model that captured the rapid change in CADSS scores with the discontinuation of esketamine. Denote the model with a linear time function as joint model 1 and the model with a linear spline time function as joint model 2.

For the time to relapse, we used the relative risk model (RRM) that depends on CADSS score trajectory at time t [redacted] (b) (4)

This model was compared to Relative Risk Model depending on CADSS score trajectory and change at time t [redacted] (b) (4)

where:

- [redacted] (b) (4)
- [redacted]
- [redacted]
- [redacted]

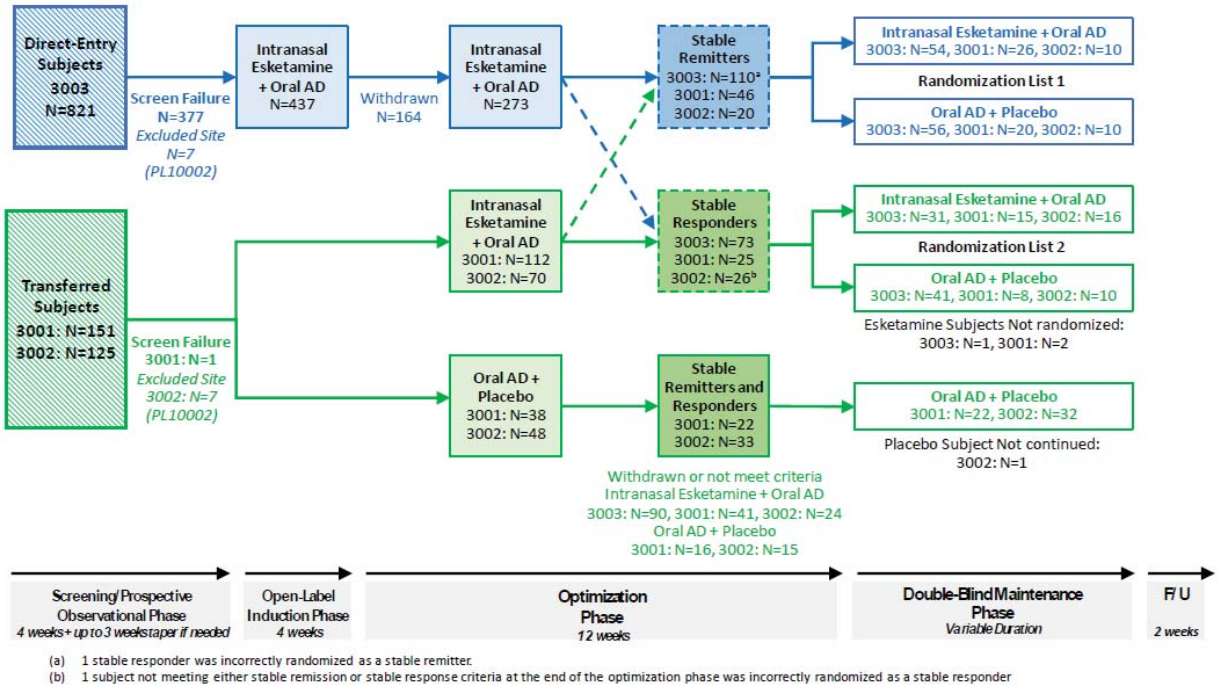
Different models were compared using Akaike Information Criteria (AIC) with the smaller AIC indicating a better fitting model.

3.2.7.3 Patient Disposition, Demographic and Baseline Characteristics

The Applicant randomized 86 patients to placebo and 90 patients to continue esketamine for a total of 175 stable remitters out of a total 705 patients starting the optimization phase, see Figure 24. Ten percent of patients withdrew from the study before termination, see Figure 25. Patients were censored if they dropped out for any reason or did not have a relapse before study termination.

Reference ID: 4392083

Figure 24: Patient Flow Diagram (TRD3003: All Subjects)



Source: Applicant's CSR TRD3003.

Figure 25: Patient Disposition (TRD3003: Maintenance Phase - Remitter Set)

	Intranasal Esk + Oral AD (N=90)	Oral AD + Intranasal Placebo (N=86)	Total (N=176)
Completed maintenance phase	82 (91.1%)	77 (89.5%)	159 (90.3%)
Subjects who relapsed	24 (26.7%)	39 (45.3%)	63 (35.8%)
Subjects who completed maintenance phase due to study termination (a)	58 (64.4%)	38 (44.2%)	96 (54.5%)
Withdrawn during maintenance phase	8 (8.9%)	9 (10.5%)	17 (9.7%)
Adverse event	1 (1.1%)	1 (1.2%)	2 (1.1%)
Pregnancy	1 (1.1%)	0	1 (0.6%)
Withdrawal by subject	3 (3.3%)	3 (3.5%)	6 (3.4%)
Other	3 (3.3%)	5 (5.8%)	8 (4.5%)

(a) Subjects were considered completers if they were relapse free at the time of the study termination

Source: Applicant's CSR TRD3003.

In the FAS, most patients were from Europe (58%). Two thirds of patients were female. Additional demographics are in Table 25.

Reference ID: 4392083

Table 25: Patient Demographic Characteristics (TRD3003: FAS – Remitter Set)

	Intranasal Esk + Oral AD (N=90)	Oral AD + Intranasal Placebo (N=86)	Total (N=176)
Age (years)			
N	90	86	176
Mean (SD)	45.4 (12.12)	46.2 (11.16)	45.8 (11.64)
Median	47.5	45.0	46.5
Range	(19; 64)	(19; 64)	(19; 64)
Age category (years), n (%)			
N	90	86	176
18-44	38 (42.2%)	37 (43.0%)	75 (42.6%)
45-64	52 (57.8%)	49 (57.0%)	101 (57.4%)
Sex, n (%)			
N	90	86	176
Male	32 (35.6%)	27 (31.4%)	59 (33.5%)
Female	58 (64.4%)	59 (68.6%)	117 (66.5%)
Race, n (%)			
N	90	86	176
Black or African American	4 (4.4%)	6 (7.0%)	10 (5.7%)
White	80 (88.9%)	76 (88.4%)	156 (88.6%)
American Indian or Alaskan native	0	1 (1.2%)	1 (0.6%)
Other	2 (2.2%)	1 (1.2%)	3 (1.7%)
Multiple	1 (1.1%)	0	1 (0.6%)
Not Reported	3 (3.3%)	2 (2.3%)	5 (2.8%)
Ethnicity, n (%)			
N	90	86	176
Hispanic or Latino	14 (15.6%)	12 (14.0%)	26 (14.8%)
Not Hispanic or Latino	73 (81.1%)	72 (83.7%)	145 (82.4%)
Not Reported	3 (3.3%)	2 (2.3%)	5 (2.8%)
Baseline weight (kg)			
N	90	86	176
Mean (SD)	82.78 (19.554)	84.21 (20.781)	83.48 (20.118)
Median	82.05	81.05	82.00
Range	(47.0; 147.0)	(44.1; 179.0)	(44.1; 179.0)
Baseline height (cm)			
N	90	86	176
Mean (SD)	169.05 (11.330)	168.60 (9.673)	168.83 (10.525)
Median	167.55	166.00	167.00
Range	(150.0; 210.3)	(150.0; 192.0)	(150.0; 210.3)
Baseline body mass index (kg/m2)			
N	90	86	176
Mean (SD)	28.9 (5.75)	29.5 (6.26)	29.2 (6.00)
Median	28.4	28.7	28.6
Range	(18; 47)	(20; 54)	(18; 54)
BMI category (kg/m2), n (%)			
N	90	86	176
Underweight<18.5	2 (2.2%)	0	2 (1.1%)
Normal 18.5-<25	19 (21.1%)	18 (20.9%)	37 (21.0%)
Overweight 25-<30	32 (35.6%)	33 (38.4%)	65 (36.9%)

Reference ID: 4392083

	Intranasal Esk + Oral AD (N=90)	Oral AD + Intranasal Placebo (N=86)	Total (N=176)
Obese 30-<40	33 (36.7%)	30 (34.9%)	63 (35.8%)
Morbidly obese >=40	4 (4.4%)	5 (5.8%)	9 (5.1%)
Employment status, n (%) (a)			
N	90	86	176
Any type of employment	57 (63.3%)	54 (62.8%)	111 (63.1%)
Any type of unemployment	23 (25.6%)	19 (22.1%)	42 (23.9%)
Other	10 (11.1%)	13 (15.1%)	23 (13.1%)
Hypertension status, n (%) (b)			
N	90	86	176
Yes	23 (25.6%)	19 (22.1%)	42 (23.9%)
No	67 (74.4%)	67 (77.9%)	134 (76.1%)
Country, n (%)			
N	90	86	176
Belgium	1 (1.1%)	1 (1.2%)	2 (1.1%)
Brazil	11 (12.2%)	11 (12.8%)	22 (12.5%)
Canada	1 (1.1%)	0	1 (0.6%)
Czech Republic	14 (15.6%)	14 (16.3%)	28 (15.9%)
France	3 (3.3%)	3 (3.5%)	6 (3.4%)
Germany	0	1 (1.2%)	1 (0.6%)
Hungary	2 (2.2%)	2 (2.3%)	4 (2.3%)
Italy	2 (2.2%)	1 (1.2%)	3 (1.7%)
Mexico	5 (5.6%)	5 (5.8%)	10 (5.7%)
Poland	19 (21.1%)	18 (20.9%)	37 (21.0%)
Slovakia	2 (2.2%)	1 (1.2%)	3 (1.7%)
Spain	2 (2.2%)	3 (3.5%)	5 (2.8%)
Sweden	3 (3.3%)	3 (3.5%)	6 (3.4%)
Turkey	4 (4.4%)	3 (3.5%)	7 (4.0%)
United States	21 (23.3%)	20 (23.3%)	41 (23.3%)
Region, n (%)			
N	90	86	176
Europe	52 (57.8%)	50 (58.1%)	102 (58.0%)
North America	22 (24.4%)	20 (23.3%)	42 (23.9%)
Other	16 (17.8%)	16 (18.6%)	32 (18.2%)
Class of oral antidepressant, n (%)			
N	90	86	176
SNRI	62 (68.9%)	58 (67.4%)	120 (68.2%)
SSRI	28 (31.1%)	28 (32.6%)	56 (31.8%)
Oral antidepressant, n (%)			
N	90	86	176
Duloxetine	47 (52.2%)	38 (44.2%)	85 (48.3%)
Escitalopram	13 (14.4%)	14 (16.3%)	27 (15.3%)
Sertraline	15 (16.7%)	14 (16.3%)	29 (16.5%)
Venlafaxine extended release (XR)	15 (16.7%)	20 (23.3%)	35 (19.9%)

Source: Applicant's TRD3003 CSR.

Reference ID: 4392083

3.2.7.4 Results and Conclusions

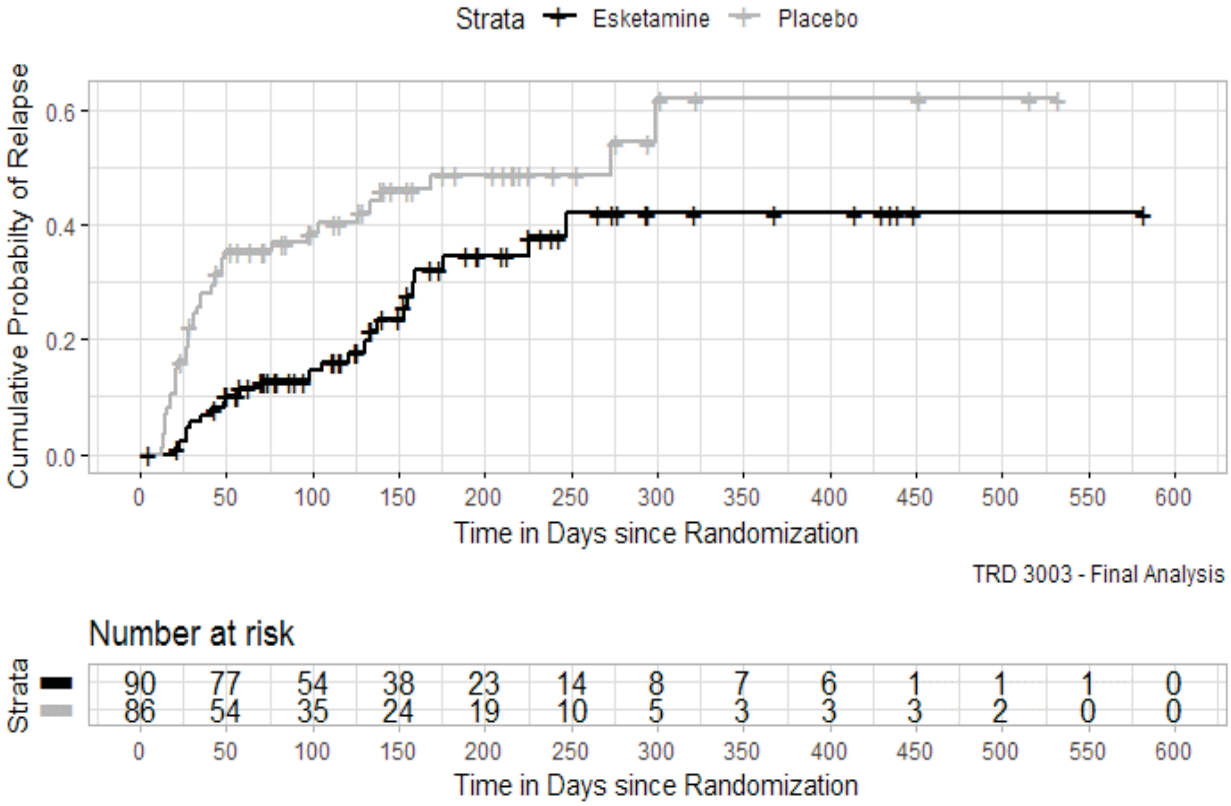
Applicant's Results

Using the pre-specified log-rank test, esketamine had a statistically significant longer time to relapse (two-sided p = 0.0029) with a median unbiased hazard ratio (HR) of 0.49 (esketamine/placebo; 95% CI: 0.29, 0.83) assuming that the hazard ratio is constant over time.

Reviewer's Notes: *However, the assumption of a constant HR did not appear to be reasonable, see Figure A4. Therefore, the hazard ratio is interpreted as the average hazard ratio over the study.*

The cumulative probability of relapse (Figure 26) rapidly increased during the first 25-30 days in the placebo arm. Forty percent of placebo patients have relapsed by 100 days after randomization. The probability of relapse increased more slowly in the esketamine arm where 40% of patients relapsed by 250 days after randomization. The MADRS total score trajectories show a similar pattern, see Figure 27 for the entire optimization and maintenance phases and Figure 28 for the last twenty days of optimization phase and first thirty days of the maintenance period. In both figures, MADRS scores start to rise early in the placebo group compared to the esketamine group. In both the placebo and esketamine arms, approximately 20% of patients survived at least 200 days after randomization without a relapse, Table 26.

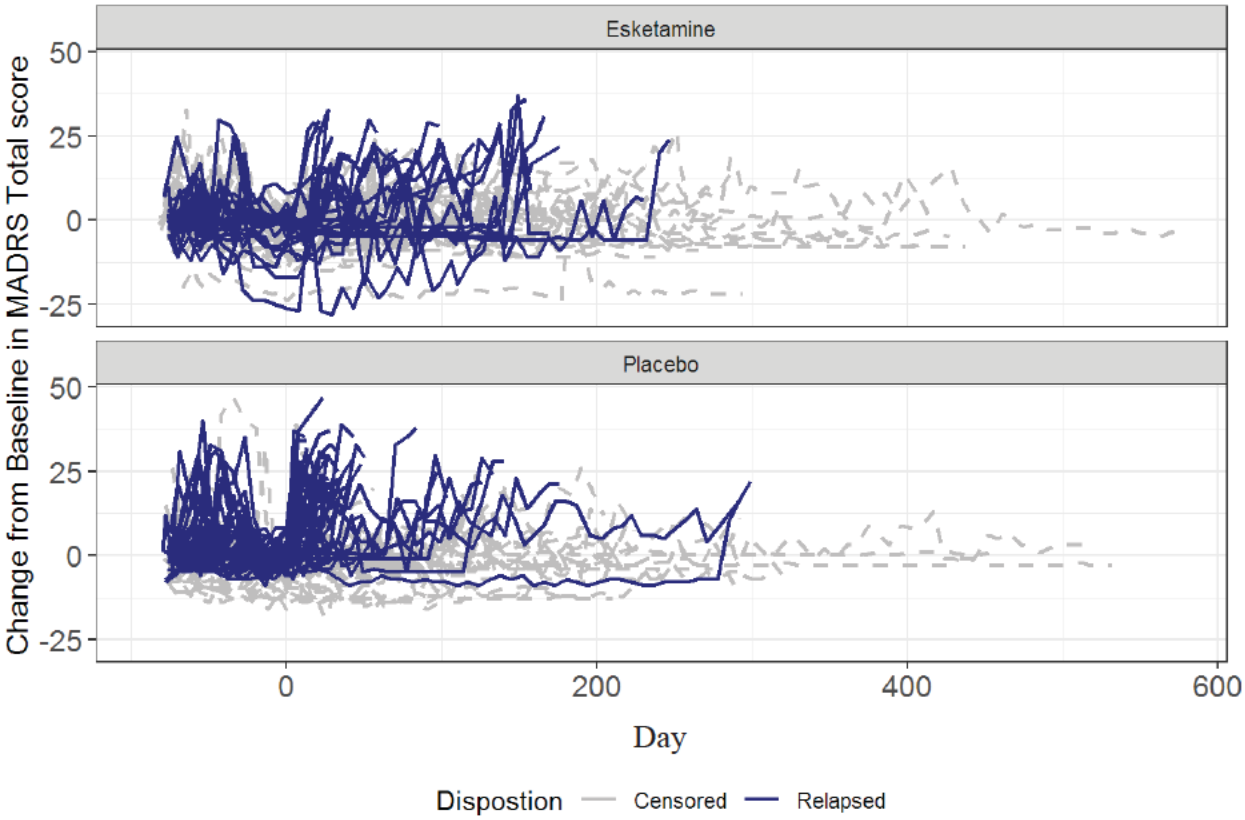
Figure 26: Time to Relapse (TRD3003: FAS - Remitter Set)



Reference ID: 4392083

Source: Reviewer.

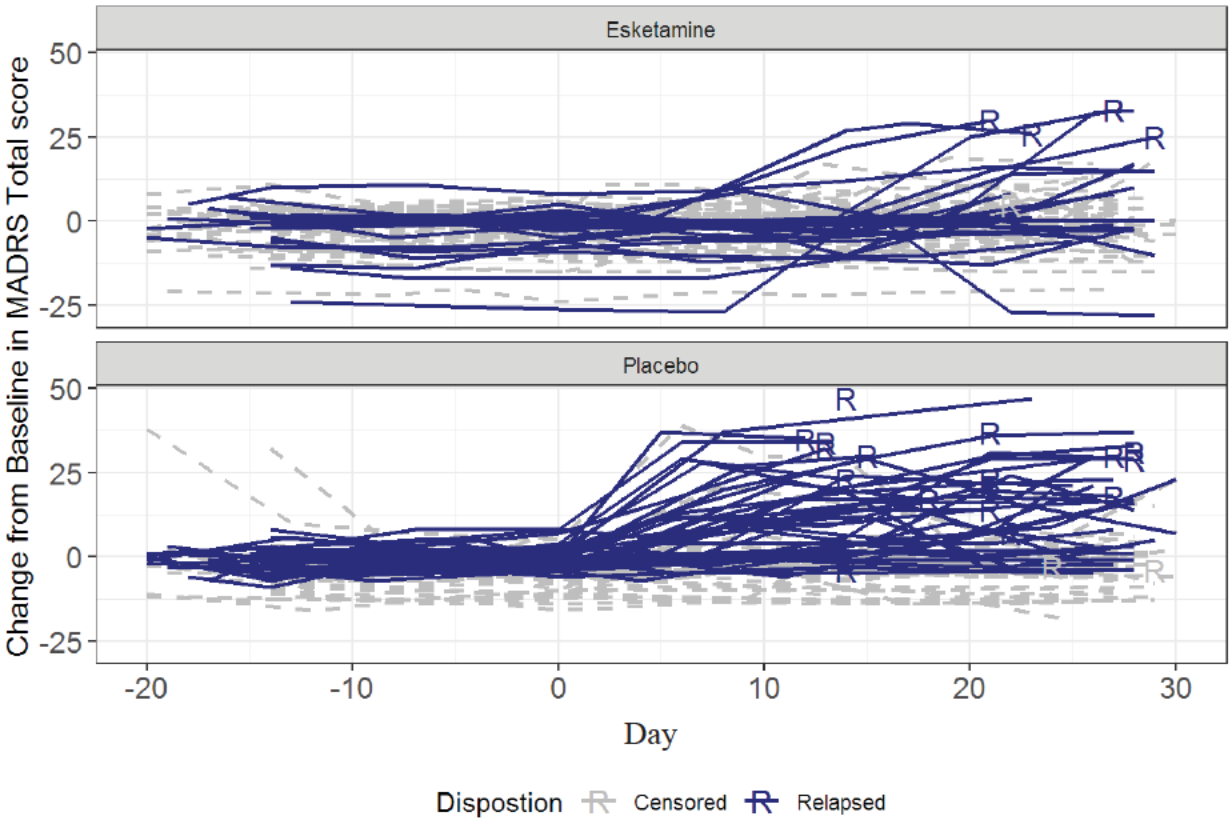
Figure 27: MADRS Total Score by Study Day - Optimization Phase Baseline (TRD3003: FAS Remitter Set)



Source: Reviewer.

Reference ID: 4392083

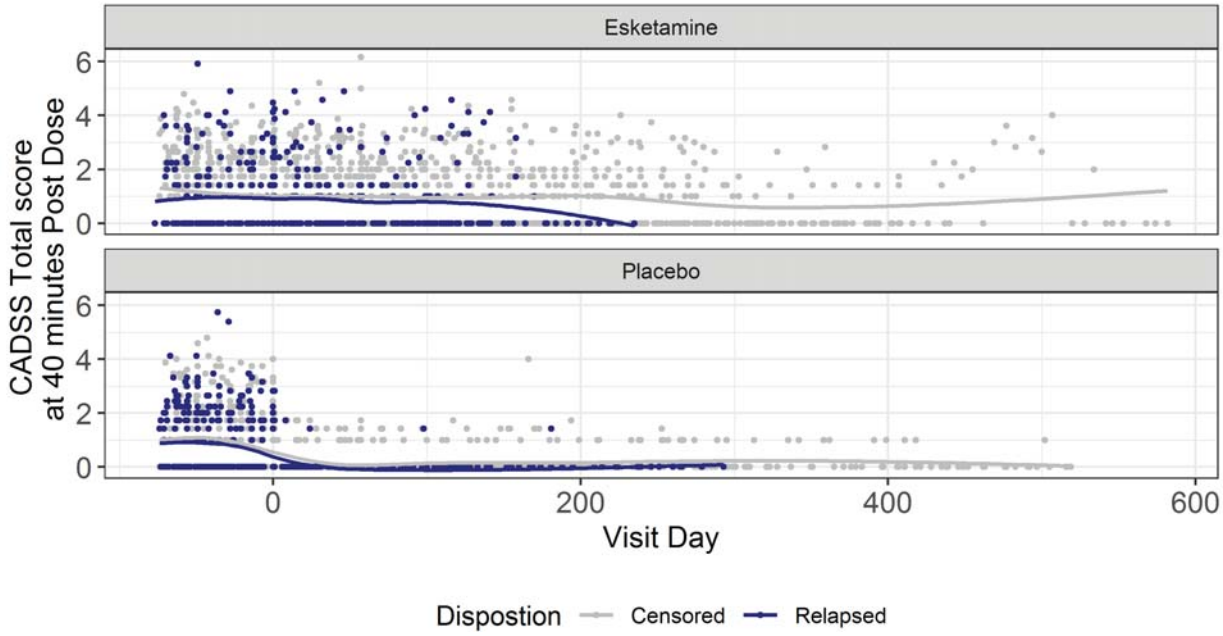
Figure 28: MADRS Total Score by Study Day - Optimization Phase Baseline - First Month of Maintenance Phase (TRD3003: FAS Remitter Set)



The average CADSS score declined in the placebo arm after the end of the optimization phase (visit day = 0), see Figure 29. This decline was present in both cohorts of patients who relapsed and who were censored. No drop in CADSS score was observed in the esketamine arm. The Applicant's post hoc sensitivity analysis assesses the effect of large changes in the CADSS score after randomization. They censored three patients who had an unspecified "clear" change in CADSS score. The hazard ratio (esketamine to placebo) changed to 0.50 (95% CI: 0.30, 0.84).

Reference ID: 4392083

Figure 29: Treatment Arm Average CADSS score Trajectories Stratified by Relapse Status (TRD3003: FAS - Remitter Set)



Source: Reviewer.

Table 26: Patients surviving at least 200 days (TRD3003: FAS - Remitter Set)

	Relapsed – N(%)	***Censored – N(%)***
Esketamine	2 (2)	21 (23)
Placebo	2 (2)	17 (20)
Denominator – Number of Patients Surviving to 200 days		
Esketamine	2 (8.7)	21 (91.3)
Placebo	2 (10.5)	17 (89.5)

Source: Reviewer.

This Reviewer fit several joint models including joint model 1 (AIC = 8910.046) and joint model 2 (AIC = 8883.019). Visual inspection of residual plots (see Appendix) showed that the normality assumption for the CADSS scores may not be reasonable. The martingale residuals do not show any systematic trends. Results of both models are reported in Table 27, Table 28, Table 29, and Table 30. Both models show a decrease in average CADSS score after randomization. In joint model 2, the esketamine HR = 0.45, an 8% change from the primary efficacy HR = 0.49. In addition, the CADSS trajectory associated with relapse (HR = 0.57 per unit increase in sqrt CADSS). The inclusion of the CADSS trajectory slopes did not improve model fit. The results are not presented in this Review.

Table 27: Longitudinal Sub-Model of Joint Model 1 (TRD3003: FAS - Remitter Set)

	Value	**SE**	**Z-value**	**P-value**

Reference ID: 4392083

(Intercept)	0.7941	0.0246	32.3214	<0.0001
Placebo	-0.8483	0.0386	-21.9656	<0.0001
Day	-0.0021	0.0002	-11.8987	<0.0001
Placebo by day	0.0026	0.0003	9.5077	<0.0001

Source: Reviewer.

Table 28: Survival Sub-Model of Joint Model 1 (TRD3003: FAS - Remitter Set)

		Value	HR	SE	Z-value	Two-sided P-value
Esketamine		-0.7530	0.4710	0.2622	-2.8722	0.0041
Association of CADSS score		-0.5252	0.5916	0.2384	-2.2034	0.0276
Baseline hazard components	log(xi.1)	-4.5955		0.2815	-16.3256	
	log(xi.2)	-4.6919		0.3190	-14.7083	
	log(xi.3)	-6.0489		0.6087	-9.9381	
	log(xi.4)	-5.8149		0.4477	-12.9897	
	log(xi.5)	-4.9574		0.3639	-13.6231	
	log(xi.6)	-6.5157		0.7249	-8.9879	
	log(xi.7)	-6.3677		0.5852	-10.8815	

Source: Reviewer.

Table 29: Longitudinal Sub-Model of Joint Model 2 (TRD3003: FAS - Remitter Set)

Longitudinal Model	Value	SE	Z-value	Two-sided P-value
(Intercept)	0.6666	0.0263	25.3798	<0.0001
Placebo	-0.6113	0.0787	-7.7635	<0.0001
Spline(day) – piece 1	-0.0023	0.0006	-3.8743	0.0001
Spline(day) – piece 2	0.0007	0.0002	3.6392	0.0003
Placebo by Spline(day) – piece 1	-0.0095	0.0057	-1.6558	0.0978
Placebo by Spline(day) – piece 2	0.0001	0.0003	0.4762	0.6340

Source: Reviewer.

Table 30: Survival Sub-Model of Joint Model 2 (TRD3003: FAS - Remitter Set)

	Survival Model	Value	HR	SE	Z-value	Two-sided

Reference ID: 4392083

						P-value
Esketamine		-0.7774	0.4510	0.2639	-2.9457	0.0032
Association of CADSS score		-0.4699	0.6251	0.2342	-2.0069	0.0448
Baseline Hazard Components	log(xi.1)	-4.6530		0.2754	-16.8926	
	log(xi.2)	-4.7268		0.3191	-14.8110	
	log(xi.3)	-6.0596		0.6121	-9.8992	
	log(xi.4)	-5.7911		0.4578	-12.6485	
	log(xi.5)	-4.8949		0.3857	-12.6913	
	log(xi.6)	-6.4048		0.7416	-8.6366	
	log(xi.7)	-6.1206		0.6100	-10.0342	

Source: Reviewer.

While the statistical significance of the CADSS trajectory is not strong (two-sided p = 0.026 – 0.045), its addition affects the point estimate of the esketamine HR (change of ~8%). These results raise concerns that dissociation affects the causal relationship between esketamine withdrawal and relapse of depression. Therefore, in consultation with Dr. Jean Kim, we consider several potential interpretations of TRD3003:

- Despite the association of dissociative symptoms with time to relapse, this potential change in patients' perception of their treatment assignment does not change the evidence that esketamine delays time to depression relapse.

- The efficacy of esketamine in delaying time to relapse depends on the subject feeling some dissociative symptoms. The subject may worsen either due to suspecting they are no longer taking active drug, or because there is some primary antidepressant effect from or association with dissociation.

The joint modeling approach combines the analysis of both the time to relapse model and CADSS trajectory model into a single model providing a direct estimate of the association between the two. However, these joint models do not consider any differences between patients entering TRD3003 from the double-blind studies (TRD3001, TRD3002) and directly entering TRD3003. There may be differences in the patient experience with open-label and double-blind esketamine during the induction and optimization periods. In addition, if the longitudinal model is miss-specified, the estimated association between time to relapse and CADSS score trajectory may be biased. In its current design, the randomized withdrawal study does not permit separating the treatment effect of esketamine into "antidepressant" and "dissociative" components.

3.3 Evaluation of Safety

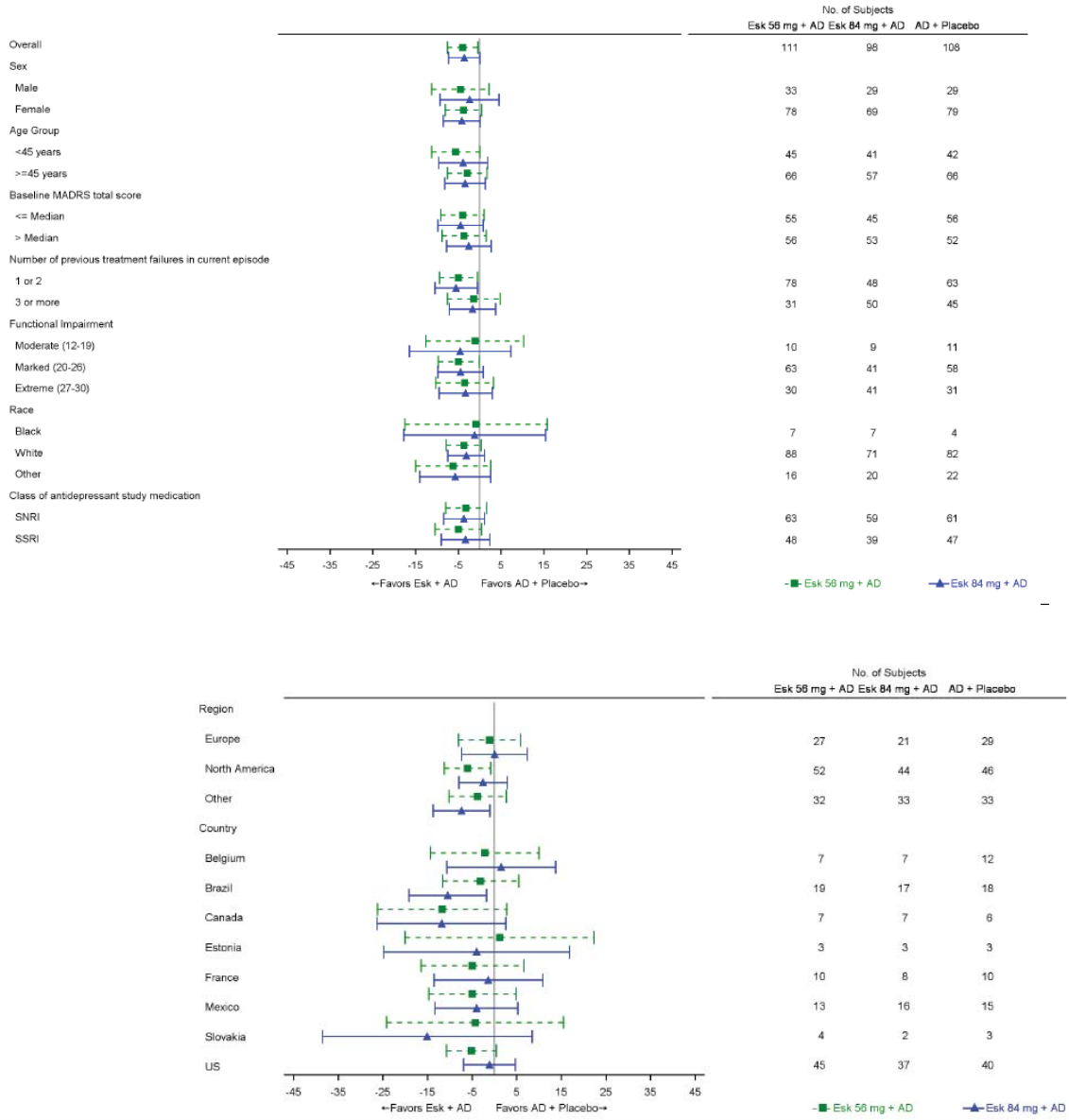
This review does not evaluate safety. Please refer to the clinical review for an evaluation of safety.

Reference ID: 4392083

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

This section contains the results of the Applicant's subgroup analyses. All subgroup analyses are post hoc, exploratory analyses and should be interpreted with care. All figures are from the Applicant's CSR for each study.

Figure 30: Subgroup Analyses (TRD3001: FAS)



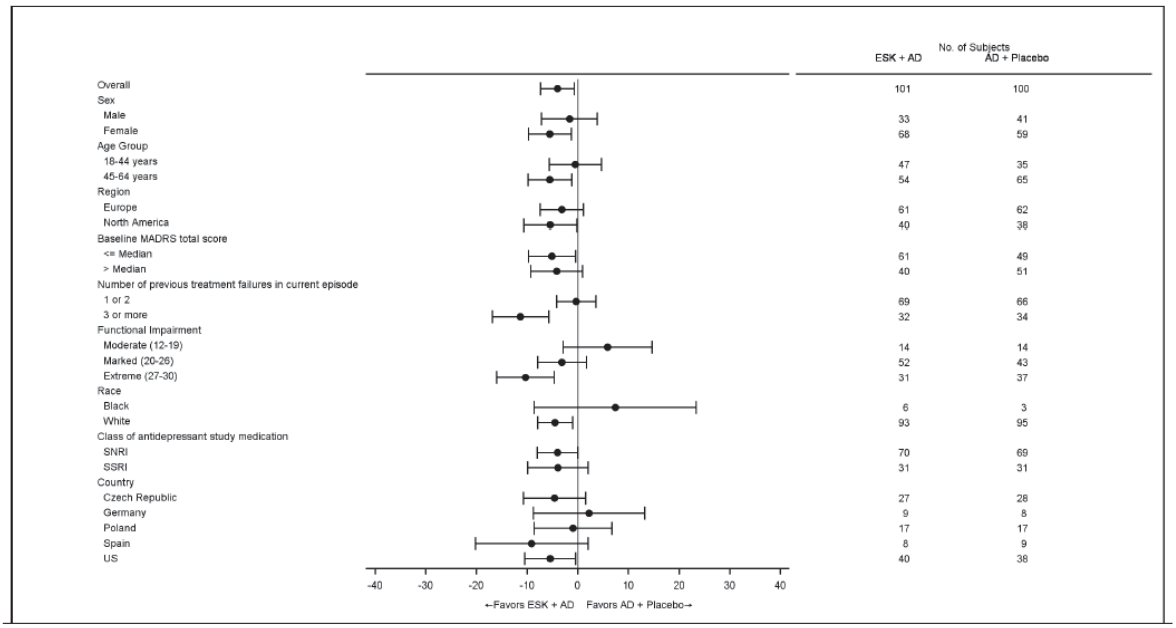
Note: Subgroups with 5 or less number of subjects are not included.

[GEFMADSG01A.RTF] [JNJ-54135419\TRD3001\DBR_FINAL2\RE_CSR\PROD\GEFMADSG01A.SAS] 21MAR2018, 17:27

Abbreviations: AD: antidepressant; Esk: esketamine; MADRS: Montgomery-Asberg depression rating sale; MMRM: mixed-effect model using repeated measures; SNRI: serotonin and norepinephrine reuptake inhibitors; SSRI: selective serotonin reuptake inhibitors; and US: United States.

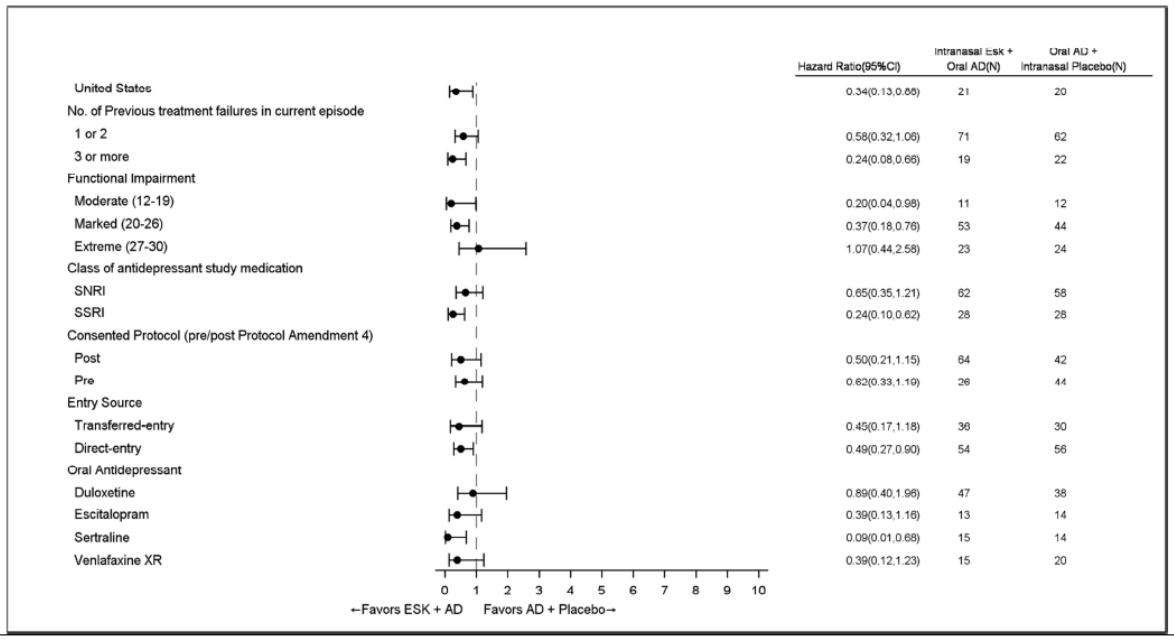
Reference ID: 4392083

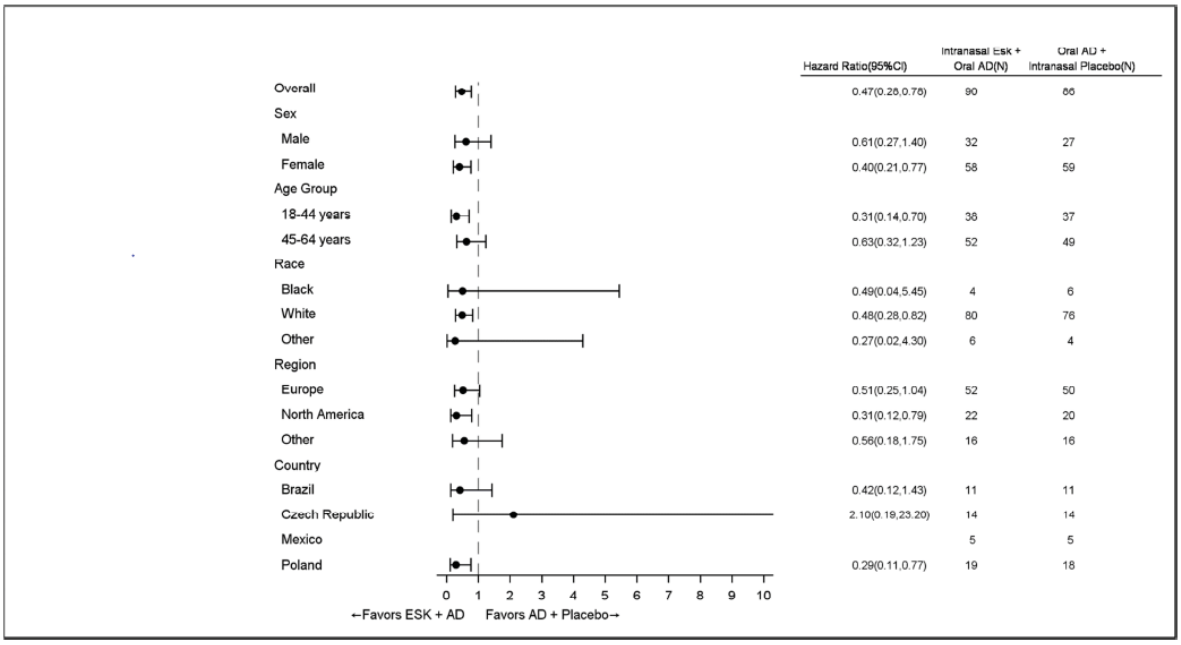
Figure 31: Subgroup Analyses (TRD3002: FAS)



Reference ID: 4392083

Figure 32: Subgroup Analyses (TRD3003: FAS - Remitter Set)





Reference ID: 4392083

In forest plots using the hazard ratio, it is considered as a summary statistics regardless of whether proportional hazards assumption holds.

4.1 Gender, Race, Age, and Geographic Region

Gender: There is no evidence for a gender subgroup difference.

Race: The small number of black patients limits any conclusions about a race effect. Whites and non-whites have similar efficacy.

Age: See elderly study TRD3005. In TRD3001 and TRD3002, there is no evidence of any efficacy difference by age.

Geographic Region: While there is variation between difference countries, there are no trends in the subgroups.

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues

This Review addressed several statistical issues:

- What esketamine doses (56 mg, 84 mg) show clinical efficacy?
- Can the relapse prevention trial ESKETINTRD3003 provide confirmatory evidence?
- What statistical evidence can be extracted from TRD 3001 after statistical testing stopped at 84 mg dose?
- Does the esketamine treatment effect change over time?

5.1.1 What esketamine doses (56 mg, 84 mg) shows clinical efficacy?

The Applicant's Phase 2 (TRD2003) and Phase 3 (TRD3001 and TRD3002) provide contradictory evidence of esketamine's dose-response. Both TRD2003 and TRD3002 suggest that higher dose (84 mg) of esketamine provided superior efficacy to lower dose (56 mg) of esketamine. However, the fixed dose study TRD3001 supports the opposite conclusion. TRD3001 suggests that 56 mg esketamine has equivalent or superior efficacy to 84 mg. While TRD3001 was the largest study (N = 342), the 84 mg dose arm had patients who failed a greater number of antidepressant therapies compared to the 56 mg dose arm in TRD3001 and both TRD2003 (N = 67) and TRD3002 (N = 223). Any dose-response information extracted from TRD3002 may be confounded by patient response to treatment. The Applicant increased esketamine dose in TRD3002 patients who had a worse clinical response and who could tolerate a higher dose. As submitted, this Application contains inconsistent dose-response information.

5.1.2 Can the relapse prevention study ESKETINTRD3003 provide confirmatory evidence?

The interpretation of TRD3003 is unclear because esketamine has acute side-effects (dissociation and sedation – not considered in this Review) that patients report. When patients are exposed to

Reference ID: 4392083

open-label esketamine and then switched to placebo, they may notice that they no longer have dissociative effects as evidenced in the decline in CADSS score after switching to placebo. However, the esketamine's treatment effect persists despite the presence of potential change in patients' perception of their treatment assignment. TRD3003 did not contain any information about whether the change CADSS score leads to change in patients' perception of their treatment assignment. While the relationship between the changes in dissociation and relapse may affect the results of TRD3003, TRD3003 provides evidence of esketamine's efficacy. Further research is needed to determine how a post-randomization side-effect potentially affects statistical inference about drug efficacy and estimate of treatment effect.

5.1.3 Additional Statistics Issues

TRD3001 suggests that the 56 mg dose arm is more efficacious than placebo in treating TRD (one-sided $p = 0.0114$); however, the 56 mg dose arm was not formally tested and would not have reached statistical significance at the original sample size of 234 patients. The initial esketamine to placebo treatment difference remained throughout the 28 days of double blind therapy observed in both TRD3001 and TRD3002.

5.2 Collective Evidence

The Applicant's four Phase 3 studies provide an inconsistent picture of esketamine's efficacy. Results from TRD3001's 56 mg dose arm and TRD3002 support esketamine's efficacy. However, the Applicant's decision to use a hierarchical, fixed sequence, multiple testing procedure stopped formal statistical testing in TRD3001 at the 84 mg dose arm. Therefore, the Applicant had to rely on TRD3003, a randomized withdrawal study, to provide evidence in a second adequate and well-controlled trial. TRD3003 provides evidence of esketamine's efficacy. However, point estimates and test statistics related to esketamine's efficacy may be influenced by the association of changes in patient's perception of acute dissociation after receiving placebo instead of esketamine.

5.3 Conclusions and Recommendations

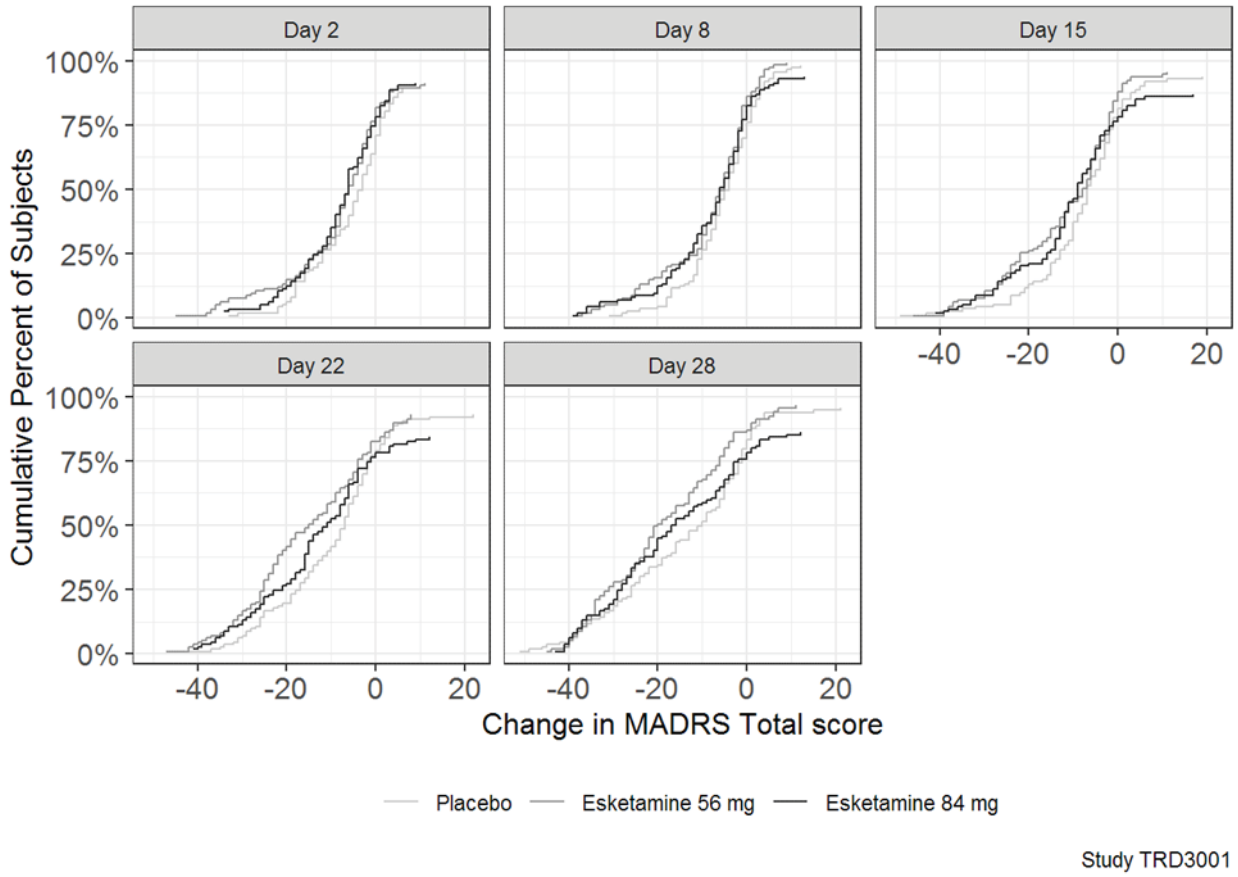
The Applicant's four submitted studies provide moderate support of esketamine's efficacy but contradictory evidence about the dose response for esketamine. Only one study, ESKETINDTRD3002, provided clear evidence for the efficacy of esketamine in treating treatment resistant depression. While ESKETINDTRD3003 demonstrated that patients randomized to placebo relapsed earlier, any interpretation of this result may be influenced by the association of time to relapse with changes in patient's dissociative symptoms. We defer to the Division of Psychiatry Products to determine if the Applicant has submitted two adequate and well controlled studies consisting of positive results in ESKETINDTRD3002 and ESKETINDTRD3003.

Reference ID: 4392083

APPENDIX

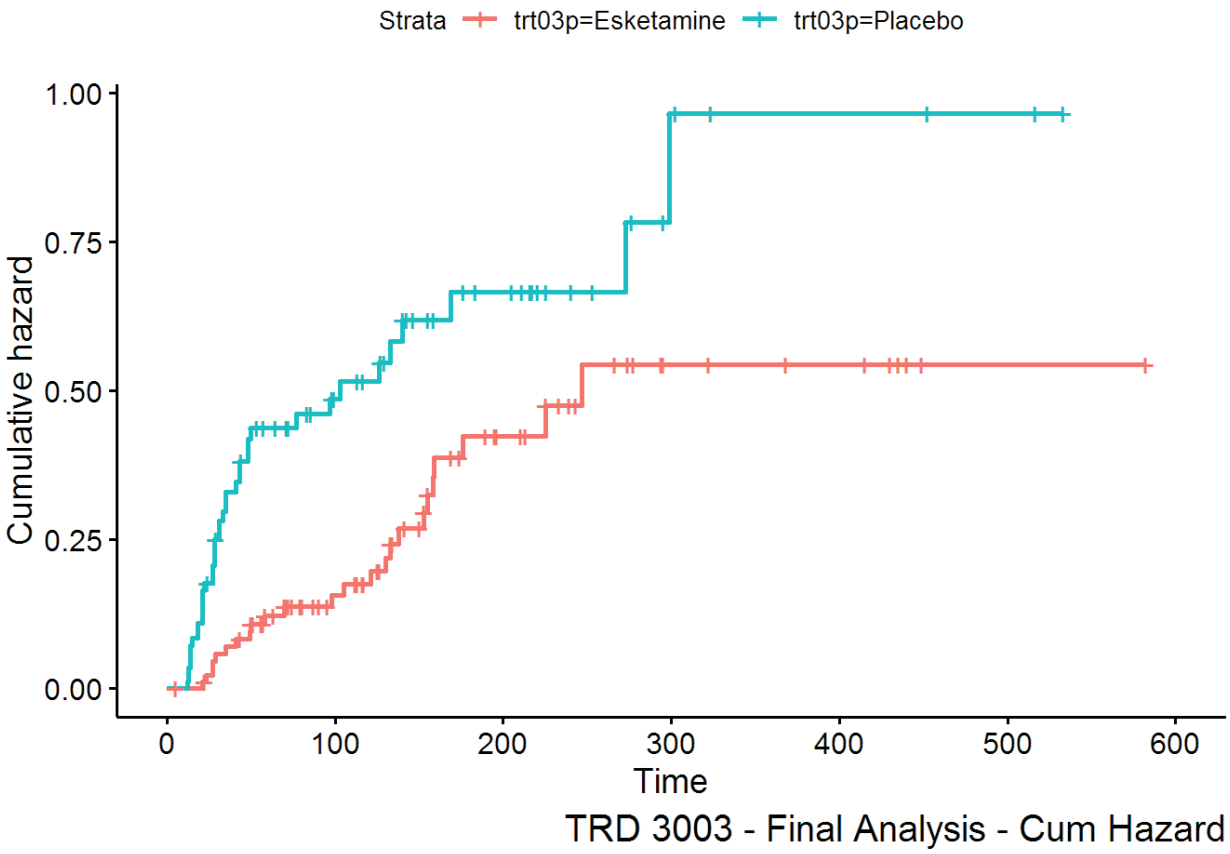
All figures in this Appendix were created by the Reviewer.

Figure A1: Empirical CDF (TRD 3001: FAS)



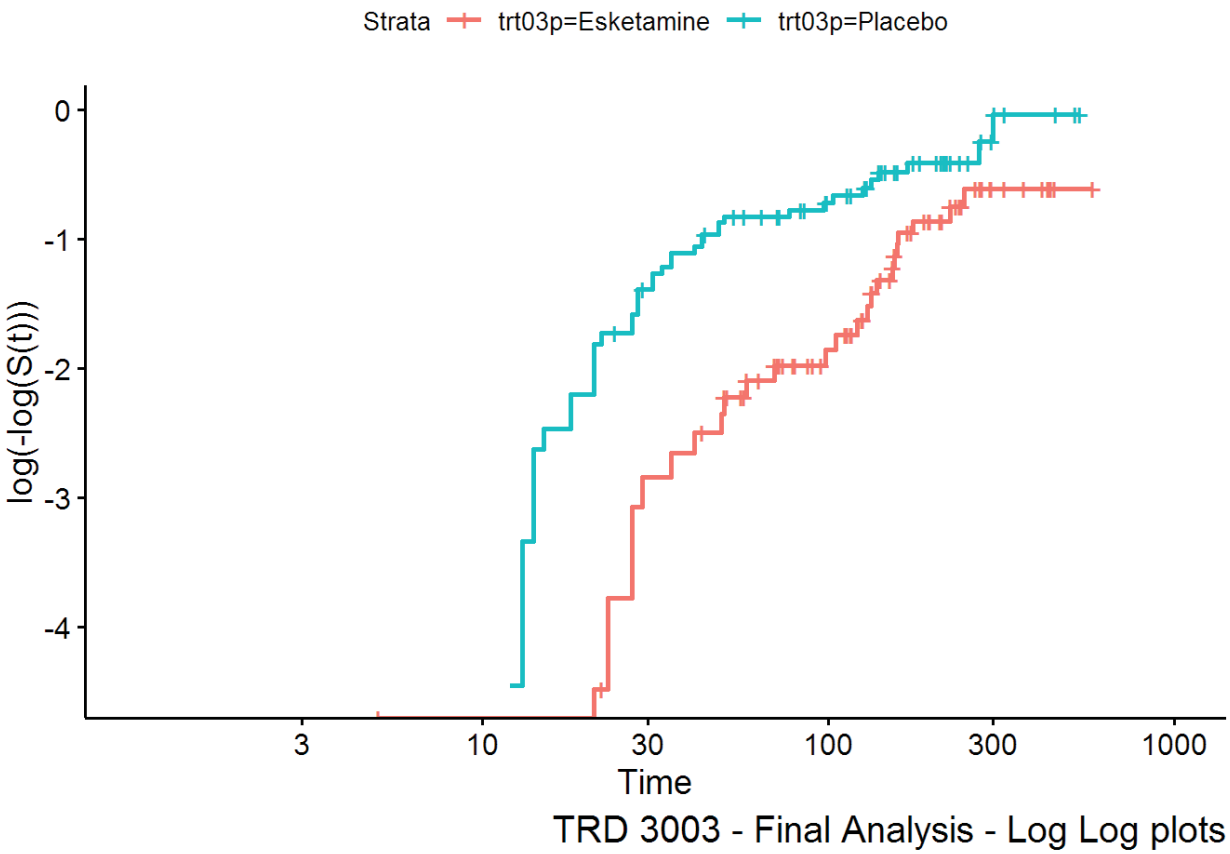
Reference ID: 4392083

Figure A2: Cumulative Hazard Function Estimates (TRD3003: FAS - Remitter Set)



Reference ID: 4392083

Figure A3: Log-log Diagnostic Plot (TRD3003: FAS – Remitter Set)



Reference ID: 4392083

Figure A4: Cumulative Event Distributions of the Time to Censoring Distribution (TRD3003: FAS – Remitter Set)

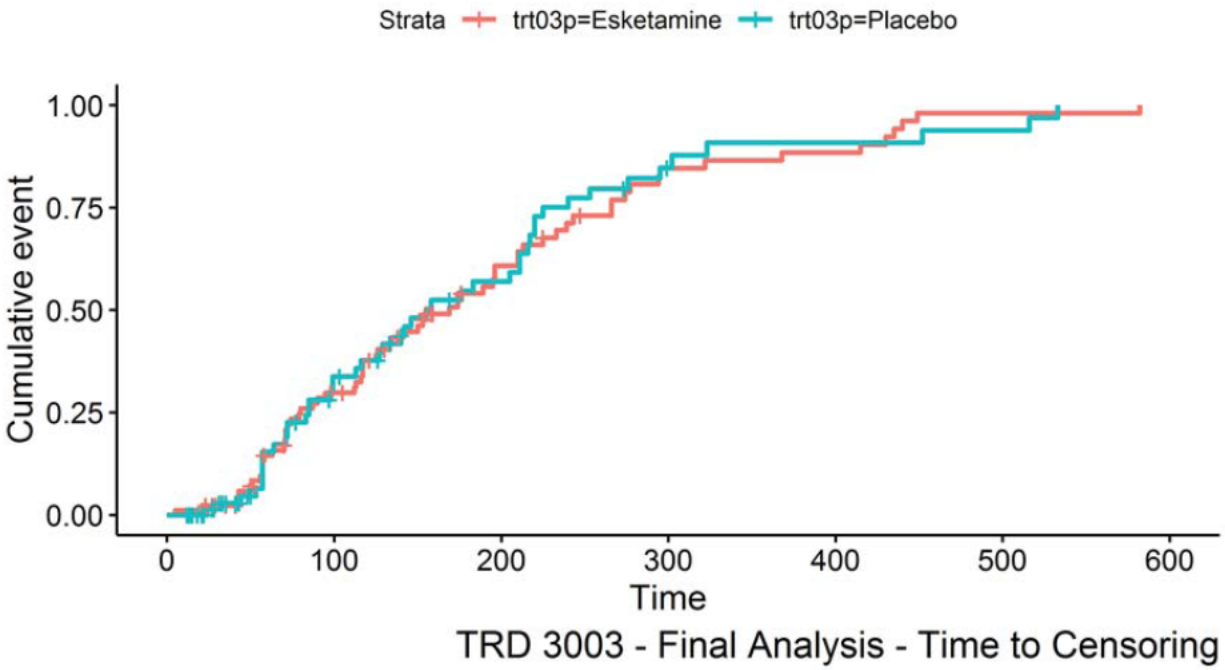
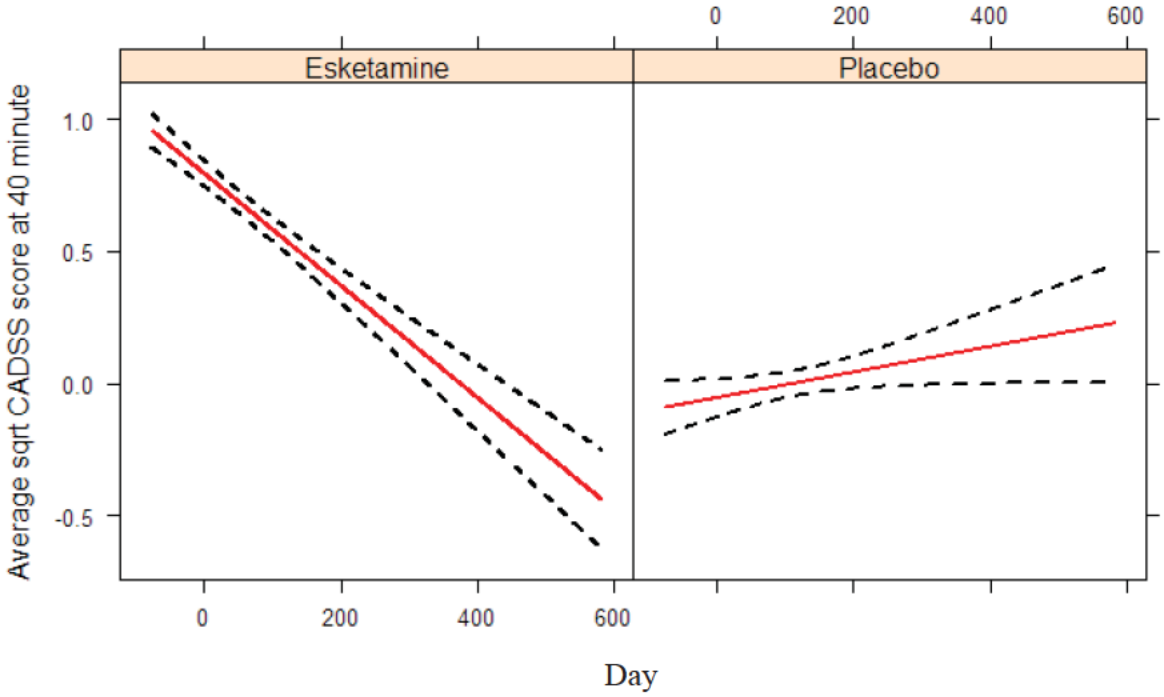


Figure A5: Predicted Average CADSS Score for Joint Model 1 (TRD3003: FAS – Remitter Set)



Reference ID: 4392083

Figure A6: Predicted Average CADSS Score for Joint Model 2 (TRD3003: FAS – Remitter Set)

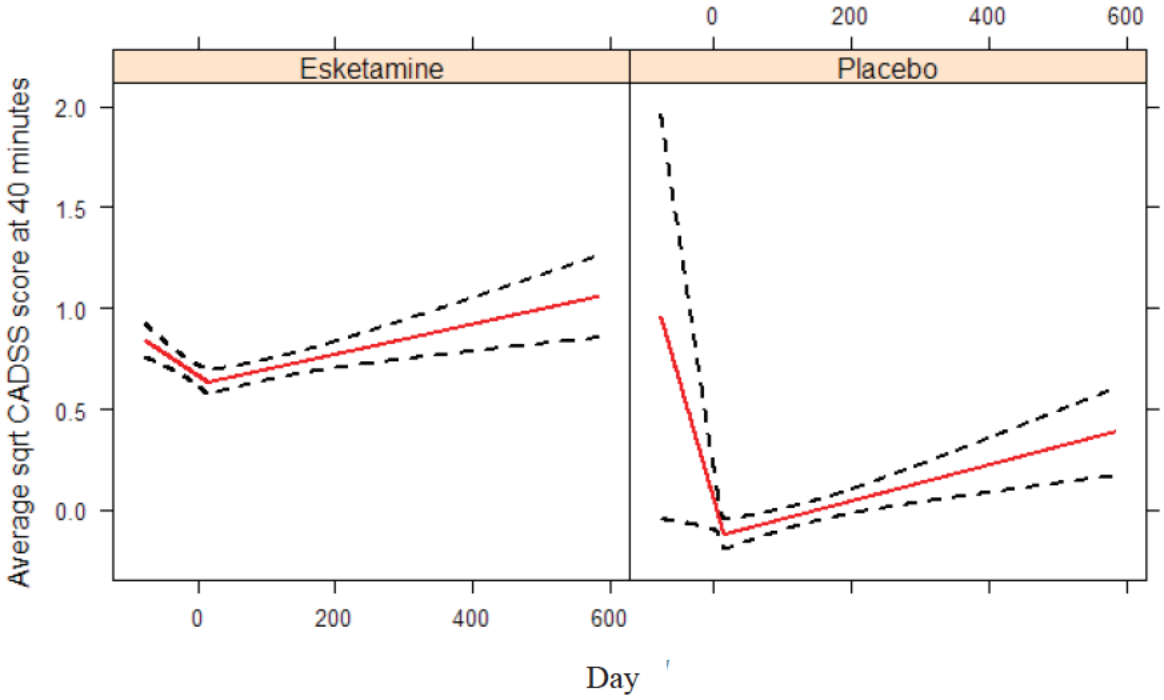
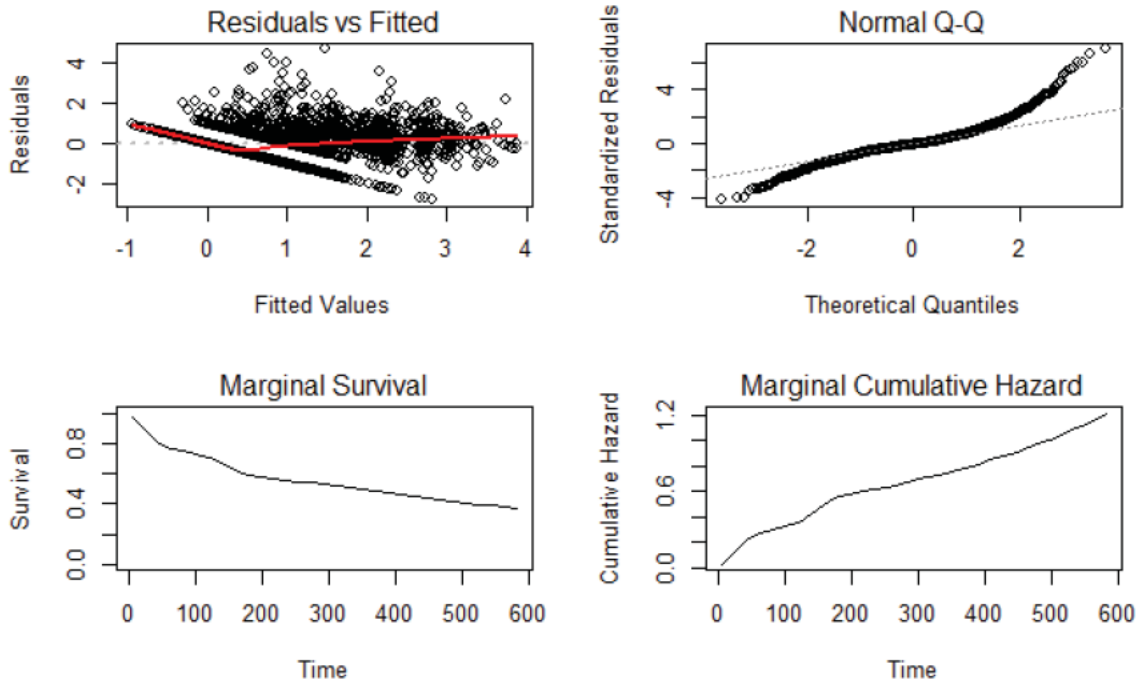


Figure A7: Residual Plots for Joint Model 1 (TRD3003: FAS - Remitter Set)



Reference ID: 4392083

Figure A8: Martingale Residuals for Joint Model 1 (TRD3003: FAS - Remitter Set)

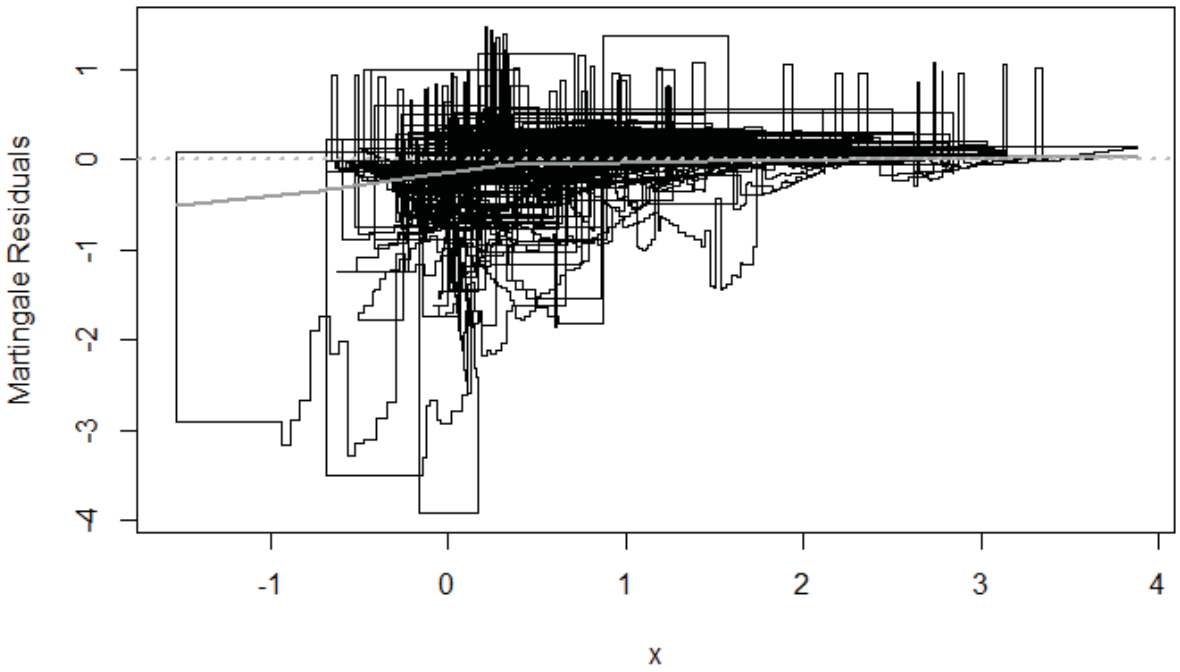
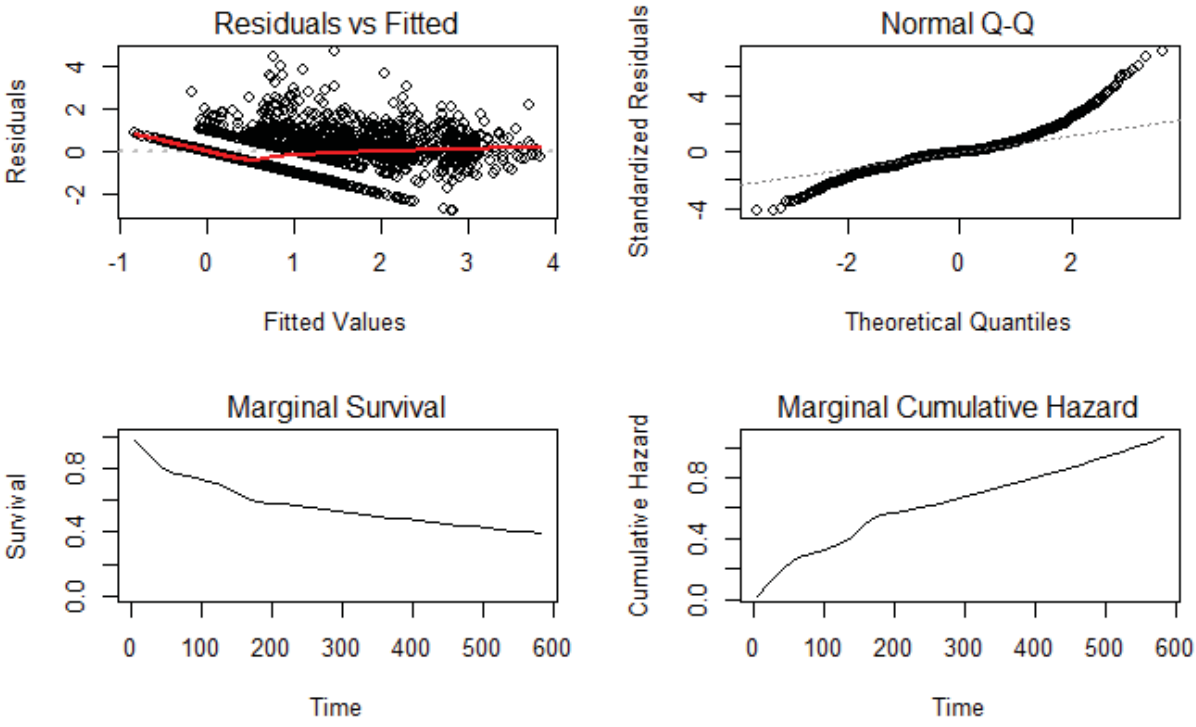
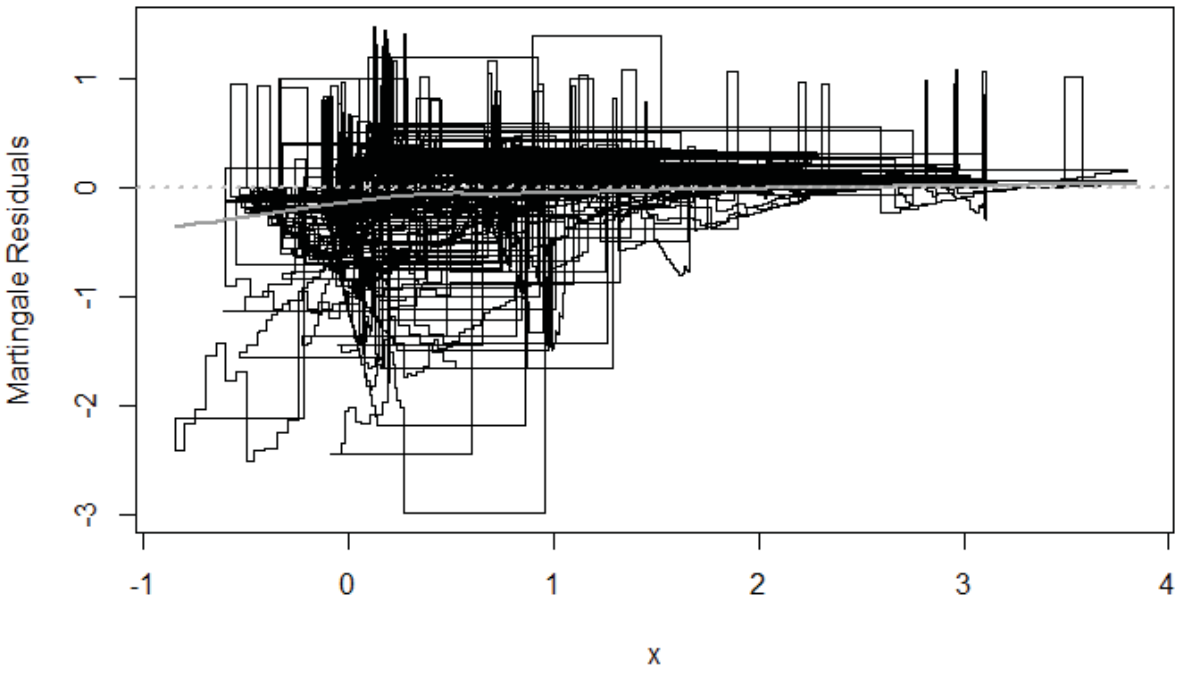


Figure A9: Residual Plots for Joint Model 2 (TRD3003: FAS - Remitter Set)



Reference ID: 4392083

Figure A10: Martingale Residuals for Joint Model 2 (TRD3003: FAS - Remitter Set)



Reference ID: 4392083

This is a representation of an electronic record that was signed electronically. Following this are manifestations of any and all electronic signatures for this electronic record.

/s/

ANDREW N POTTER
02/15/2019 04:25:50 PM

PEILING YANG
02/15/2019 09:16:35 PM

HSIEN MING J HUNG
02/15/2019 09:43:08 PM

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Science
Office of Biostatistics

Statistical Review and Evaluation

CLINICAL STUDIES

NDA/Serial Number:	211243
Drug Name:	JNJ-54135419 ([Esketamine, S-(+)-ketamine])
Indication:	Treatment of patients with treatment-resistant depression (TRD)
Study number:	54135419TRD1015
Applicant:	Janssen Research & Development, LLC
Date(s):	Date of Document: 7/15/2018 Completion date: 1/2/2019
Review Priority:	S
Biometrics Division:	DB VI
Statistical Reviewer:	Qianyu Dang, Ph.D., Lead Mathematical Statistician, CSS supporting team/DBVI/OB
Concurring Reviewers:	Yi Tsong, Ph.D., Division Director, DBVI/OB
Medical Division:	Controlled Substance Staff
The CSS Team:	Jovita Randall-Thompson, Ph.D., Pharmacologist, OD/CSS Dominic Chiapperino, Ph.D., Acting Director, OD/CSS
Project Manager:	Sandra Saltz, OD/CSS

Keywords: Crossover design; Human abuse potential study; Self-reported endpoint;

Reference ID: 4389574

TABLE OF CONTENTS

LIST OF TABLES ***3***

LIST OF FIGURES ***3***

1. EXECUTIVE SUMMARY ***4***

2. REVIEW REPORT ON STUDY 54135419TRD1015 ..... ***6***

2.1. OVERVIEW *6*

2.1.1. Objectives of the study 6

2.1.2. Study design 6

2.1.4. Number of Subjects 7

2.1.5. Abuse potential Endpoints 8

2.1.6. Statistical methodologies used in the Sponsor's analyses 8

2.3. SUMMARY OF SPONSOR REPORTED ANALYSIS RESULTS *9*

2.4. REVIEWER'S ASSESSMENT *10*

2.4.1. Primary Analysis 10

2.4.1.1. Descriptive statistics 10

2.4.1.2. Statistical Testing 12

2.4.2. Secondary Analysis 13

2.4.2.1. Descriptive Statistics for Take Drug Again and Overall Drug Liking 14

2.4.2.2. Descriptive Statistics for Hallucinating, Floating, Detached and Spaced out 17

2.4.2.2. Statistical Testing of Secondary Outcomes 25

3. CONCLUSION ***28***

Reference ID: 4389574

List of Tables

Table 1: Summary statistics for Drug Liking Emax (N=34) 10
Table 2: Primary analysis results on Drug Liking Emax (N=34) 13
Table 3: Summary statistics for Take Drug Again Emax and Overall Drug Liking Emax (N=34) 15
Table 4: Summary statistics for Emax Hallucinating, Floating, Detached and Spaced out (N=34) 20
Table 5: Statistical analysis results on Take Drug Again Emax (N=34) 26
Table 6: Statistical analysis results on Overall Drug Liking Emax (N=34) 26
Table 7: Statistical analysis results on Hallucinating Emax (N=34) 26
Table 8: Statistical analysis results on Floating Emax (N=34) 27
Table 9: Statistical analysis results on Detached Emax (N=34) 27
Table 10: Statistical analysis results on Spaced Out Emax (N=34) 27

List of Figures

Figure 1: Heat Map by Treatment for Drug Liking Emax 11
Figure 2: The mean time course profiles in 12 hours on Drug Liking VAS by treatment (N=34) 12
Figure 3: Heat map by treatment for Take Drug Again Emax 14
Figure 4: Heat map by treatment for Overall Drug Liking Emax 15
Figure 5: Bar Plot for mean responses of Take Drug Again VAS at Hours 8 (N=34) 16
Figure 6: Bar Plot for mean responses of Overall Drug Liking VAS at Hours 8 VAS (N=34) 16
Figure 7: Heat map by treatment for Hallucinating Emax 17
Figure 8: Heat map by treatment for Floating Emax 18
Figure 9: Heat map by treatment for Detached Emax 19
Figure 10: Heat map by treatment for Spaced Out Emax 20
Figure 11: The mean time course profiles in 12 hours on Hallucinating VAS by treatment (N=34) 22
Figure 12: The mean time course profiles in 12 hours on Floating VAS by treatment (N=34) 23
Figure 13: The mean time course profiles in 12 hours on Detached VAS by treatment (N=34) 24
Figure 14: The mean time course profiles in 12 hours on Spaced Out VAS by treatment (N=34) 25

Reference ID: 4389574

1. Executive Summary

The applicant, Janssen Research & Development, LLC., submitted the results from the human abuse potential study 54135419TRD1015 (refer as study 1015) for the assessment of abuse potential of intranasal esketamine.

Study 1015 was a single-center, single-dose, double-blind, double-dummy, placebo-controlled, randomized crossover Study. The primary objective was to evaluate the abuse potential of esketamine in nondependent, recreational polydrug users of perception-altering drugs. Forty-one subjects were randomized to the Treatment Phase. Of these, 34 subjects completed the study.

The primary endpoint was Drug Liking Emax. The secondary endpoints included in the reviewer's secondary analysis were Take Drug again, Overall Drug liking, Hallucinating, Floating, Detached and Spaced out.

The sponsor's analysis did not use completer's population defined in 2017 FDA guidance of HAP studies. The reviewer's statistical analysis was based on 34 subjects who completed all periods though such difference did not change the conclusions.

Like IV racemic ketamine, both two doses of intranasal esketamine showed large increases in mean of VAS scores immediately after dosing with peak effects occurring by 20 minutes post-dose. The mean VAS time profile of intranasal esketamine was also similar to IV racemic ketamine.

The results from the reviewer's primary analysis on Emax Drug Liking show that

- IV racemic ketamine had significantly higher maximum liking than placebo by the margin of 15, which validated the study;
- there was no significant difference in maximum liking between IN esketamine 84 mg and IV racemic ketamine;
- there was no significant difference in maximum liking between IN esketamine 112 mg and IV racemic ketamine;

The results from the secondary analysis on Take Drug again, Overall Drug liking, Hallucinating, Floating, Detached and Spaced out indicate that:

- there was no significant difference in maximum VAS score between IN esketamine 84 mg and IV racemic ketamine in Take Drug again, Overall Drug liking, Hallucinating, Floating, Detached and Spaced out;
- there was no significant difference in maximum VAS score between IN esketamine 112 mg and IV racemic ketamine in Take Drug again and Overall Drug liking. However, IN esketamine 112 mg had significantly higher maximum VAS score than IV racemic ketamine in Hallucinating, Floating, Detached and Spaced out;

Reference ID: 4389574

In conclusion, the measure of Emax Drug Liking, Take Drug again and Overall Drug liking after IN esketamine 84 mg and 112 mg dosing are comparable with IV racemic ketamine. However, IN esketamine 112 mg had significantly higher maximum VAS score than IV racemic ketamine in Hallucinating, Floating, Detached and Spaced out, indicating possible higher perceptual effects with high dose of esketamine.

Reference ID: 4389574

2. Review report on Study 54135419TRD1015

2.1. Overview

Study CP130-1011 was was a single-center, single-dose, double-blind, double-dummy, placebo-controlled, randomized crossover Study in nondependent, recreational users of perception-altering drugs.

2.1.1. Objectives of the study

Primary Objective:

- The primary objective of this study was to evaluate the abuse potential of esketamine in nondependent, recreational polydrug users of perception-altering drugs.

Secondary Objective:
To evaluate

- relationship between select measures of abuse potential and the dose of intranasal esketamine;
- safety and tolerability of intranasal administered esketamine and intravenous (IV) racemic ketamine;
- pharmacokinetics (PK) of intranasal esketamine and IV racemic ketamine.

2.1.2. Study design

The study consisted of a Screening Phase and a Treatment Phase (TP).

Screening Phase
After providing written informed consent, subjects were enrolled in the Screening Phase consisting of a Screening Visit and a Qualification Session (QS). The Screening Visit determined eligibility for participation and included a review of the inclusion and exclusion criteria. Eligibility was further evaluated during the QS.

On Day -1 of the Qualification Session, subjects were admitted to the study center. Those who still met the enrollment criteria were randomized to receive the following:

Treatment Sequences and Study Drugs Administered*		
Sequence	Day 1	Day 2
	TRT A	**TRT B**
1	IV placebo	0.5 mg/kg IV racemic ketamine
	IN placebo	IN placebo
	TRT B	**TRT A**
2	0.5 mg/kg IV racemic ketamine	IV placebo
	IN placebo	IN placebo

Legend: *Study drugs administered concurrently
TRT = Treatment; IV = intravenous; IN = intranasal

The order (Sequence 1 or 2) was specified by a randomization schedule and study drugs were

Reference ID: 4389574

administered in a double-blind manner.

Subjects were eligible to continue into the TP if they fulfilled the following 4 criteria:

- Able to complete the study-related procedures, and could tolerate the IV racemic ketamine dose and intranasal placebo;
- Exhibited a ≥15 point difference, relative to placebo, in maximum response (Emax) on Drug Liking at the Moment within the first 4 hours after administration of 0.5 mg/kg of IV racemic ketamine over 40 minutes;
- Exhibited an acceptable placebo response, defined as ≥40 and ≤60 points for Drug Liking at the Moment during the 4 hours after administration of placebo; and
- Exhibited consistency as determined by the investigator of subjective effect reports on VAS for High, Good Effects, and Any Effects.

TP

The TP was a randomized, single-dose, double-blind, double-dummy, placebo-controlled phase in which 4 treatments (A, B, C, and D) were administered in a cross-over manner (ie, 1 treatment in each period) to measure the likelihood of abuse. Eligible subjects were readmitted to the study center on Day -1 of each of the 4 periods of the TP for baseline assessments.

Treatment Sequences and Study Drugs Administered*				
Sequence	Period 1	Period 2	Period 3	Period 4
3	**TRT A** IV placebo IN placebo	**TRT D** IV placebo IN 112 mg esketamine	**TRT B** 0.5 mg/kg IV racemic ketamine IN placebo	**TRT C** IV placebo IN 84 mg esketamine
4	**TRT B** 0.5 mg/kg IV racemic ketamine IN placebo	**TRT A** IV placebo IN placebo	**TRT C** IV placebo IN 84 mg esketamine	**TRT D** IV placebo IN 112 mg esketamine
5	**TRT C** IV placebo IN 84 mg esketamine	**TRT B** 0.5 mg/kg IV racemic ketamine IN placebo	**TRT D** IV placebo IN 112 mg esketamine	**TRT A** IV placebo IN placebo
6	**TRT D** IV placebo IN 112 mg esketamine	**TRT C** IV placebo IN 84 mg esketamine	**TRT A** IV placebo IN placebo	**TRT B** 0.5 mg/kg IV racemic ketamine IN placebo

*Study drugs administered concurrently on Day 1 of each period.
TRT: Treatment; IV: intravenous; IN: intranasal

On Day -1 of Period 1 of the TP the investigator or designated personnel provided to the subjects an in-depth review the pharmacodynamic VAS assessments to be used. On Day 1 and Day 2 of the TP at prespecified time points listed in the Time and Events Schedule, the subjects performed the serial pharmacodynamic assessments for abuse potential.

Reviewer's comments: This review is for PD outcomes. Therefore, from now on any information, analysis, results and conclusion are based on PD outcomes of TP.

2.1.4. Number of Subjects

A total of 55 subjects participated in the Qualifying Session (QS): 27 subjects (Sequence AB) and 28 subjects (Sequence BA) (Table 6). Of the 55 subjects, 41 subjects completed the QS and 14 subjects were withdrawn: 12 subjects (6 subjects each in Sequence AB and

Reference ID: 4389574

Sequence BA) did not meet all pre-defined criteria and 2 subjects (Sequence AB) were withdrawn due to "Other" reasons (1 subject was dosed outside of the protocol specified dosing window and the other subject was unable to complete Treatment 1 within the protocol defined window).
A total of 41 subjects participated in the Treatment Phase (TP): 10 subjects each in Sequence ADBC, Sequence CBDA, and Sequence DCAB and 11 subjects in Sequence BACD. Of the 41 subjects, 34 subjects completed the TP and 7 subjects were withdrawn: 1 subject (Sequence ADBC) due to an AE, and 6 due to 'Other' reason (4 subjects [1 each in Sequence ADBC and Sequence BACD and 2 in Sequence DCAB] due to PI discretion, 1 subject [Sequence BACD] due to being dosed outside of the protocol specified dosing window, and 1 subject [Sequence BACD] due to a positive urine drug screen at Period 3 for cocaine and benzodiazepines).

Reviewer's comments: The sponsor's analysis did not use completer's population defined in 2017 FDA guidance of HAP studies. The reviewer's statistical analysis was based on 34 subjects who completed all treatment periods.

2.1.5. Abuse potential Endpoints

Pharmacodynamics
A series of subjective effect reports on visual analog scale (VAS) were performed to assess abuse potential to register respondents' answers to a closed-end question along a continuum gradient. The primary dependent measure was the VAS Drug Liking at the Moment (bipolar).
The tests to be performed were classified as follows:

VAS for Balancing Measures	Positive Effects at the Moment	Negative Effects at the Moment	Perceptual Effects	Others
Drug Liking at the Moment (bipolar)	VAS for High (unipolar)	VAS for Bad (unipolar)	VAS for Hallucinating (unipolar)	VAS for Dizziness (unipolar)
Overall Drug Liking (bipolar)	VAS for Good (unipolar)	VAS for Nausea (unipolar)	VAS for Floating (unipolar)	VAS for Any Effects (unipolar)
Take Drug Again (bipolar)			VAS for Detached (unipolar)	VAS for Sedation (unipolar)
			VAS for Spaced Out (unipolar)	Drug Similarity
			VAS for Vision Clear, Crisp (bipolar)	

The primary endpoint is Drug Liking Emax.

2.1.6. Statistical methodologies used in the Sponsor's analyses

Descriptive statistics were calculated for the pharmacodynamic measures, which were organized by pharmacologic effect (ie, balancing measures, positive effects, negative effects, perceptual, and others).Values and change from baseline at each time point and maximum (E_{max}) or minimum (E_{min}) response (as appropriate) of the post-dose measurement were included.
Plots were prepared using individual and mean scores for each of the pharmacodynamic tests (excluding Overall Drug Liking, Take Drug Again, and Drug Similarity) versus nominal times. Similar plots were prepared based on individual and mean change from baseline. In addition, box plots were generated at each time point and based on peak response for each measure (excluding

Reference ID: 4389574

Overall Drug Liking, Take Drug Again, and Drug Similarity) with treatment groups displayed side-by-side. For each measure, E_{max} or E_{min} scores (as appropriate) were analyzed using a mixed-effects model, with treatment, period, and sequence as fixed effects and subject nested within sequence as a random effect; baseline (predose) measurement were used as a covariate, as applicable, (excluding Overall Drug Liking, Take Drug Again, and Drug Similarity). The pairwise comparisons were between IV racemic ketamine and placebo and each dose of intranasal esketamine and placebo. For all pairwise comparisons, the point estimates, associated 95% confidence intervals (CI), and p-values were derived from the mixed-effects model. All statistical tests will be 2-sided with the significance level set at 0.05. To be included in the analysis, a subject considered evaluable if she or he had received all scheduled doses or had participation terminated, but had previously received a placebo

The residuals from the mixed-effects model were investigated for centrality of distribution. Measures that were not centrally distributed were investigated for centrality of natural log-transformed data. Measures that did not fit either condition were analyzed nonparametrically using the Kruskal–Wallis test; where there was an overall treatment effect, pairwise comparisons were conducted using the Wilcoxon signed-rank test.

To establish model sensitivity, pair-wise comparisons between IV racemic ketamine and placebo were performed. Assay sensitivity was assumed if IV racemic ketamine separates from placebo for VAS Drug Liking at the Moment at 1 or more postdose time points.

Reviewer's comments: We do not recommend any data transformation. In the case of non-normality in the residual of mixed model, paired t-test and Sign test will be used after examining the assumptions for each test.

2.2. Data Location

The datasets used in the reviewer's analysis are located at

\\cdsesub1\evsprod\nda210730\0001\m5\datasets\cp130-1011\analysis\adam\datasets\

2.3. Summary of Sponsor Reported Analysis Results

The primary measure for Drug Liking at the Moment was assessed by a 100-point bipolar VAS. Greater responses for IV ketamine (in the direction of liking>50) relative to placebo (scores at midpoint 50) are readily apparent for assessments conducted during the initial 1-hour post dose interval. The mean of the maximum responses was 84.24. Mean differences in LS means between IV ketamine and placebo were statistically significant ($p<0.0001$) from 10 minutes to 1 hour post dose, inclusive, confirming assay sensitivity.

The time course shows that, overall, the mean Drug Liking at the Moment scores for esketamine were higher than placebo, particularly within 1 hour after administration. The mean maximum responses for the 84-mg and 112-mg doses of esketamine were 83.52 and 84.64, respectively, and were statistically significantly greater than the mean maximum response for placebo (50.53) ($p<0.0001$).

IV ketamine and intranasal esketamine had significant effects compared to placebo on secondary endpoints commonly used to assess subjective drug effects that measure balance of effects, positive, negative, sedative and perceptual drug effects.

Reference ID: 4389574

Furthermore, a similar rank order was observed for the 3 active treatments with regards to median values of reported drug similarity. The subjects perceived them as being most similar to a dissociative drug or ketamine.

Reviewer's comments: the sponsor did not follow the statistical testing procedure in 2017 FDA guidance for HAP studies.

2.4. Reviewer's Assessment

2.4.1. Primary Analysis

2.4.1.1. Descriptive statistics

Table 1 summarizes the mean, standard deviation (SD), minimum (Min), the first quartile (Q1), median (Med), the third quartile (Q3), and maximum (Max) for the 6 treatments in the study for the primary endpoint Drug Liking Emax.

Table 1: Summary statistics for Drug Liking Emax (N=34)

TRT	Mean	SD	Min	Q1	Med	Q3	Max
Esketamine 112 mg	83.7	15.0	50	72	88	95	100
Esketamine 84 mg	82.7	13.0	54	75	82	94	100
Racemic ketamine	83.6	15.5	50	72	86	100	100
Placebo	50.4	1.1	50	50	50	50	56

As summarized in Table 1, there was no much difference in means of maximum liking between IN Esketamine 112 mg, Esketamine 84 mg and IV racemic ketamine. Figure 1 is the heat map by treatment for the primary endpoint.

Reference ID: 4389574

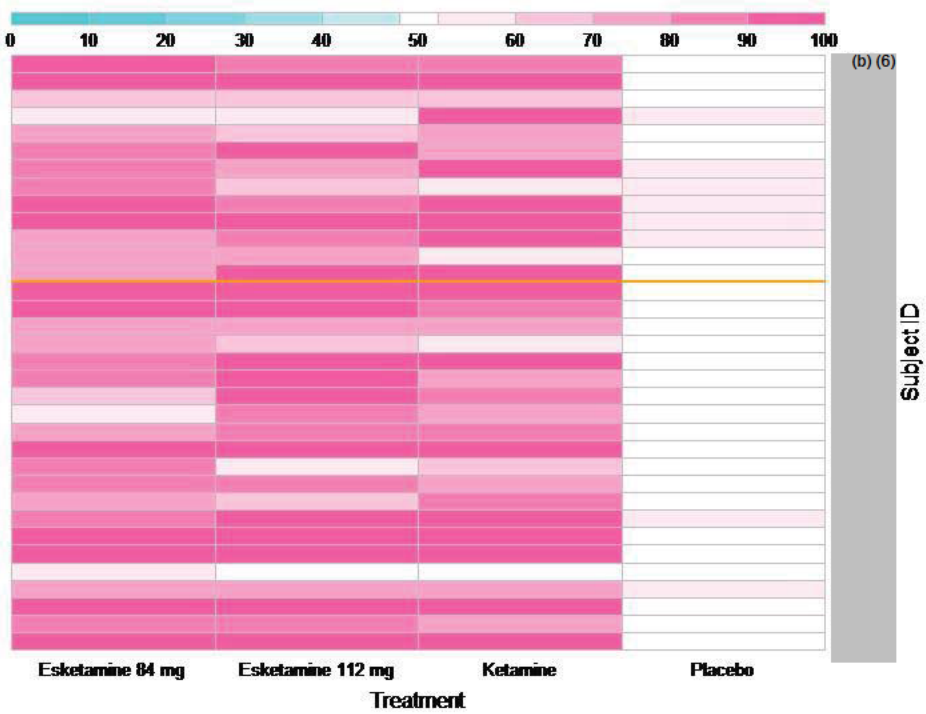


Figure 1: Heat Map by Treatment for Drug Liking Emax

The heat map presents maximum liking in color for each subject by each treatment. The orange line on the heat map separates females from males.

Figure 2 is the mean time course profiles by treatment for Drug Liking VAS. Data were collected at hours 0.17, 0.33, 0.67, 1.0, 1.5, 2.0, 2.5, 3.0, 4.0, 8.0, and 12.0. The peak mean responses for IN esketamine 112 mg, esketamine 84 mg and IV racemic ketamine are 73.4, 76.2 and 78.9, respectively. All 3 treatments showed large increases in VAS scores with peak effects occurring by 20 minutes post dose. The peak mean of drug liking under 3 treatments dropped quickly within 2 hour postdose.

Reference ID: 4389574

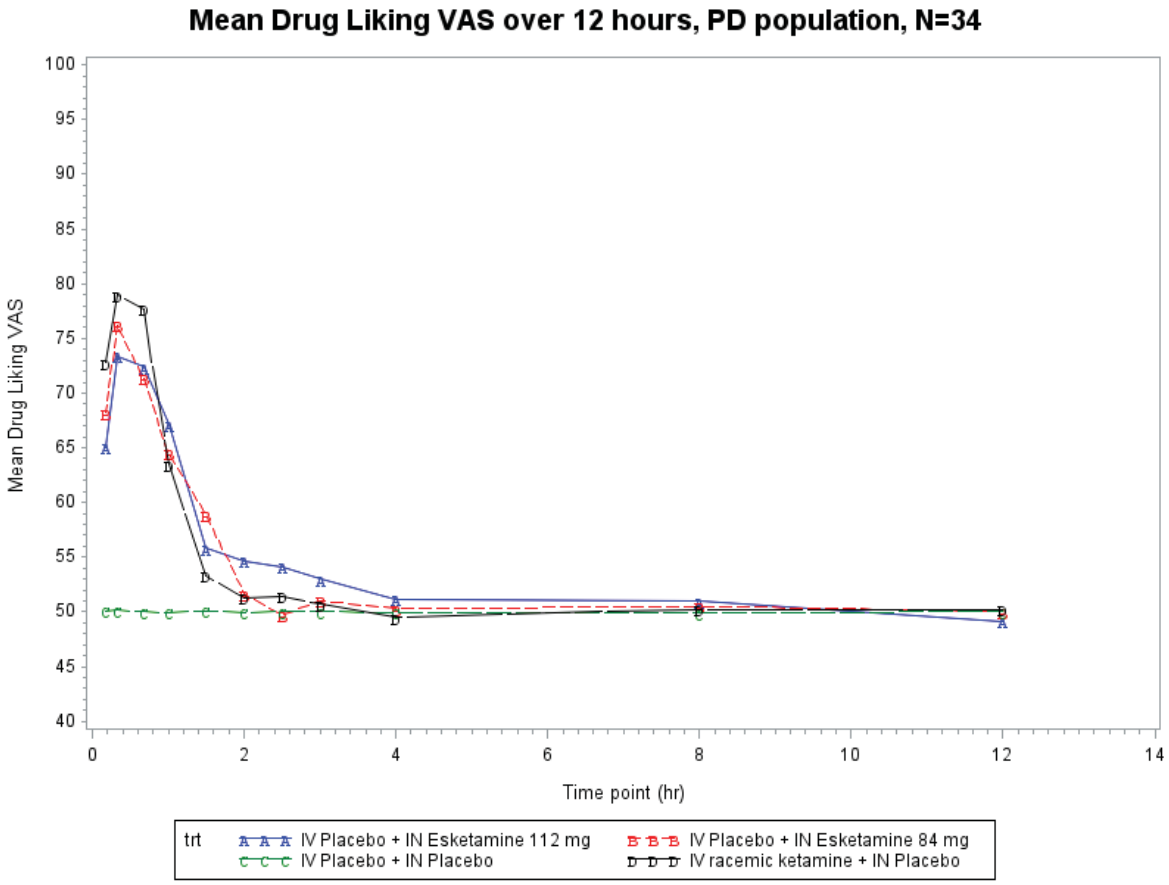


Figure 2: The mean time course profiles in 12 hours on Drug Liking VAS by treatment (N=34)

2.4.1.2. Statistical Testing

To evaluate abuse potential of IN esketamine, the following comparisons were performed for the primary endpoint, Drug Liking Emax.

1. IV racemic ketamine versus P
2. IN esketamine 112 mg versus IV racemic ketamine
3. IN esketamine 84 mg versus IV racemic ketamine
4. IN esketamine 112 mg versus P (only if H_0 of test 2 was rejected)
5. IN esketamine 84 mg versus P (only if H_0 of test 3 was rejected)

The comparisons #1 were for the study validation. If the comparison of a dose of IN esketamine did not have a statistically significantly lower mean than IV racemic ketamine, the comparison between the dose of IN esketamine and placebo would not be performed.

The statistical model used in the reviewer's primary analysis was a mixed-effects model which included sequence, period and treatment as fixed effects, and subject as a random effect. With

Reference ID: 4389574

heteroscedasticity adjustment, the Shapiro-Wilk W-test on the residual was not statistically significant for Drug Liking VAS with a p-value greater than 0.01.
The FDA 2017 Guidance recommends the following hypotheses:

1. $H_0 : \mu_C - \mu_P \le \delta_1$ versus $H_a : \mu_C - \mu_P > \delta_1$,
2. $H_0 : \mu_C - \mu_T \le \delta_2$ versus $H_a : \mu_C - \mu_T > \delta_2$, and
3. $H_0 : \mu_T - \mu_P \ge \delta_3$ versus $H_a : \mu_T - \mu_P < \delta_3$,

where *C*, *T* and *P* denote IV racemic ketamine, IN esketamine and placebo. The sponsor did not pre-specify δs. The δ_1=15, δ_2=0 and δ_3=11 were used in the reviewer's analysis.
The statistical analysis results are listed in Table 3. Since null hypothesis in test 2 for both doses of IN esketamine were not rejected, test 3 were not performed.

Table 2: Primary analysis results on Drug Liking Emax (N=34)

Pairwise Comparison	Mean Diff /Med Diff	StdErr /IQR	p-value	95% CI	
				LCL	UCL
Racemic ketamine -P	33.2	2.6	<.001	28.8	Infty
Racemic ketamine - Esketamine 112 mg	-0.088	2.1	0.52	-3.64	Infty
Racemic ketamine - Esketamine 84 mg	1.23	3.35	0.35	-4.27	Infty

The reviewer's primary analysis showed that for Drug Liking Emax:

- IV racemic ketamine had means statistically significantly greater than placebo, that validated the study;
- IN esketamine 112 mg and 84 mg both had no statistically different mean VAS score in Emax Drug Liking from IV racemic ketamine.

2.4.2. Secondary Analysis

Because this was a human abuse potential study, secondary endpoints used in the reviewer's analysis were Take Drug Again, Overall Drug Liking, Hallucinating, Floating, Detached and Spaced out.

Reference ID: 4389574

2.4.2.1. Descriptive Statistics for Take Drug Again and Overall Drug Liking

Figures 3 and 4 are the heat maps by treatment for Take Drug again Emax and Overall Drug liking Emax.

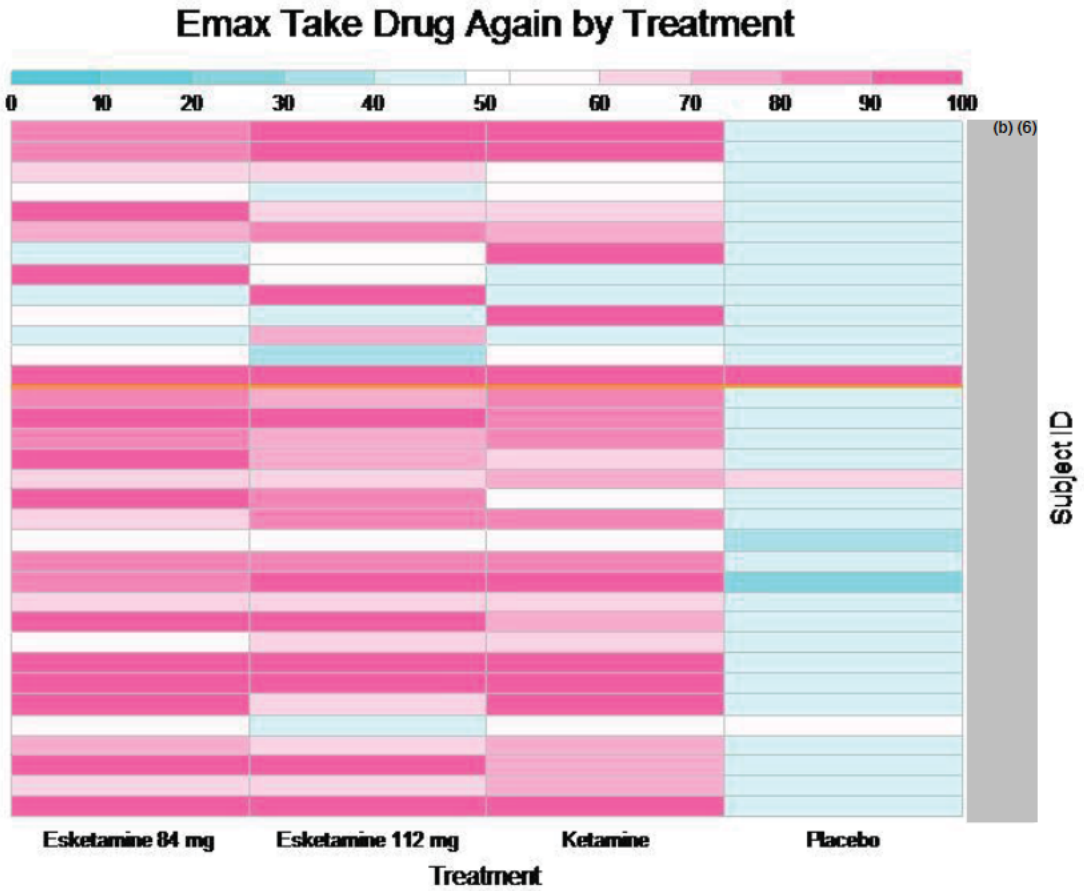


Figure 3: Heat map by treatment for Take Drug Again Emax

Reference ID: 4389574

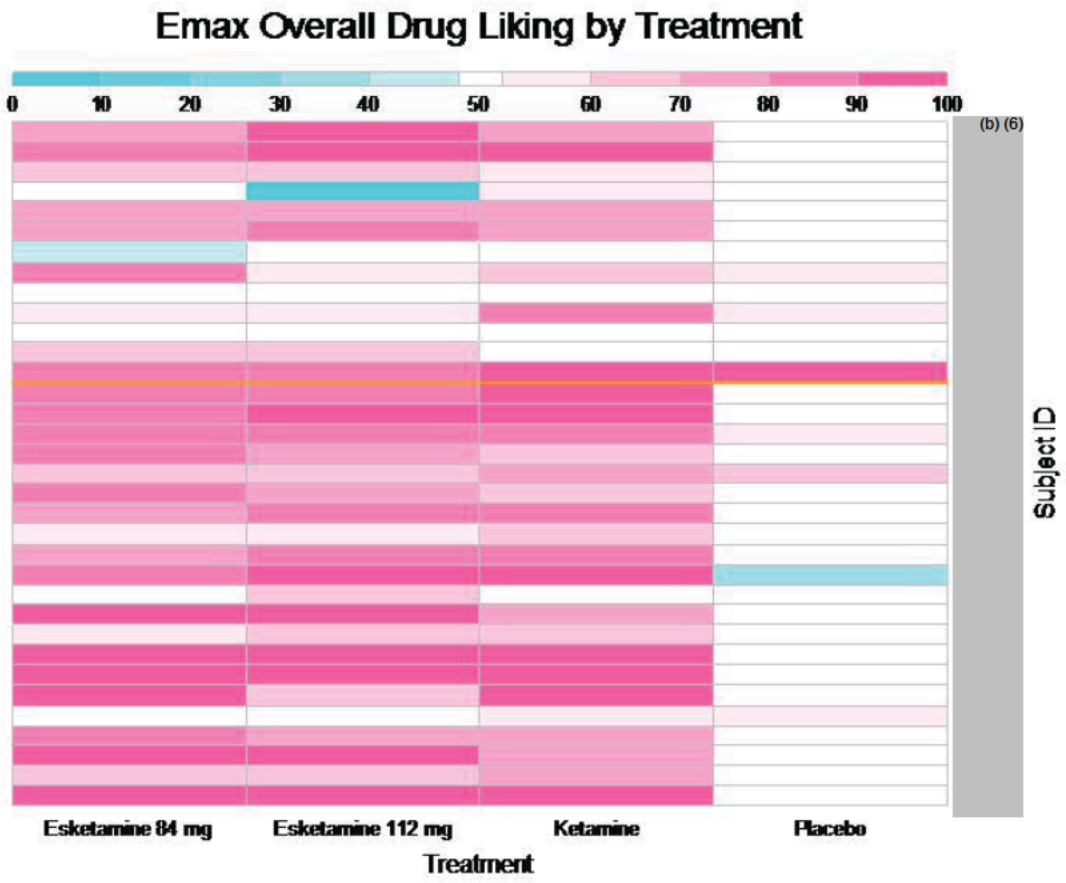


Figure 4: Heat map by treatment for Overall Drug Liking Emax

Table 3 summarizes the mean, standard deviation (SD), minimum (Min), the first quartile (Q1), median (Med), the third quartile (Q3), and maximum (Max) for the 4 treatments in the study for the secondary endpoints of Take Drug Again Emax and Overall Drug Liking Emax.

Table 3: Summary statistics for Take Drug Again Emax and Overall Drug Liking Emax (N=34)

Measure	TRT	Mean	SD	Min	Q1	Med	Q3	Max
TDA VAS N=34	Esketamine 112 mg	76.6	19.9	32	63	77	98	100
	Esketamine 84 mg	77.1	18.5	46	60	83.5	92	100
	Racemic ketamine	76.9	17.7	48	60	75	93	100
	Placebo	50.8	10.0	29	50	50	50	100
ODL VAS N=34	Esketamine 112 mg	74.3	20.7	8	61	78	91	100

Reference ID: 4389574

	Esketamine 84 mg	74.9	17.2	49	56	78.5	88	100
	Racemic ketamine	75.2	17.3	50	61	76.5	91	100
	Placebo	51.4	8.1	36	50	50	50	92

Figure 5 and Figure 6 are bar plots for mean responses of Take Drug Again VAS and Overall Drug Liking VAS, respectively. From these graphs, one may see all three groups of IN esketamine 112 mg, 84 mg and IV racemic ketamine had similar mean scores.

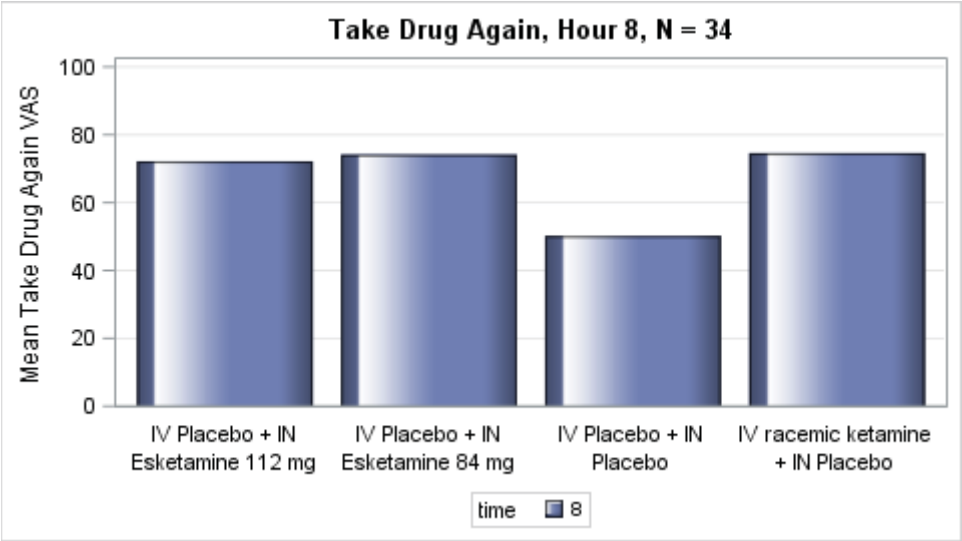


Figure 5: Bar Plot for mean responses of Take Drug Again VAS at Hours 8 (N=34)

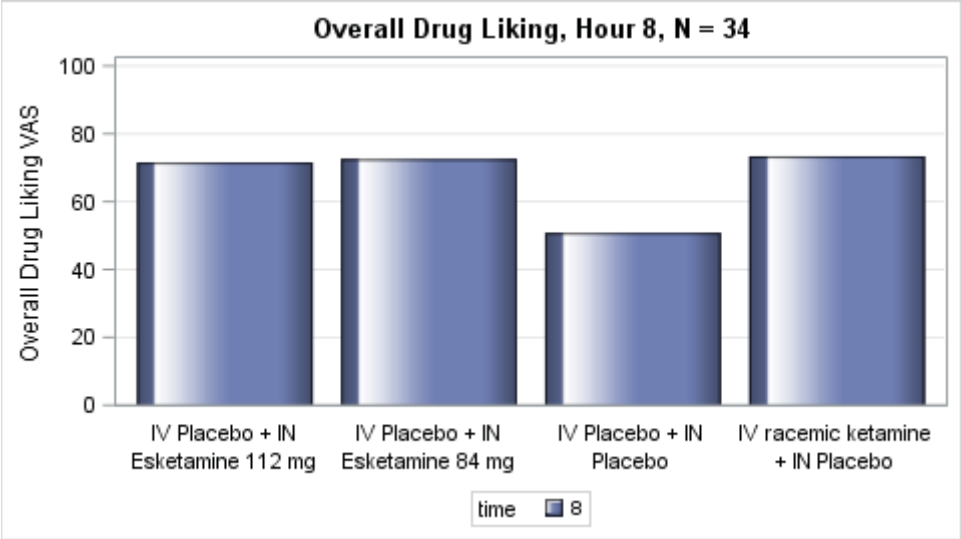


Figure 6: Bar Plot for mean responses of Overall Drug Liking VAS at Hours 8 VAS (N=34)

Reference ID: 4389574

2.4.2.2. Descriptive Statistics for Hallucinating, Floating, Detached and Spaced out

Figure 7, Figure 8, Figure 9 and Figure 10 are the heat maps by treatment for perceptual effects of Emax Hallucinating, Floating, Detached and Spaced out.

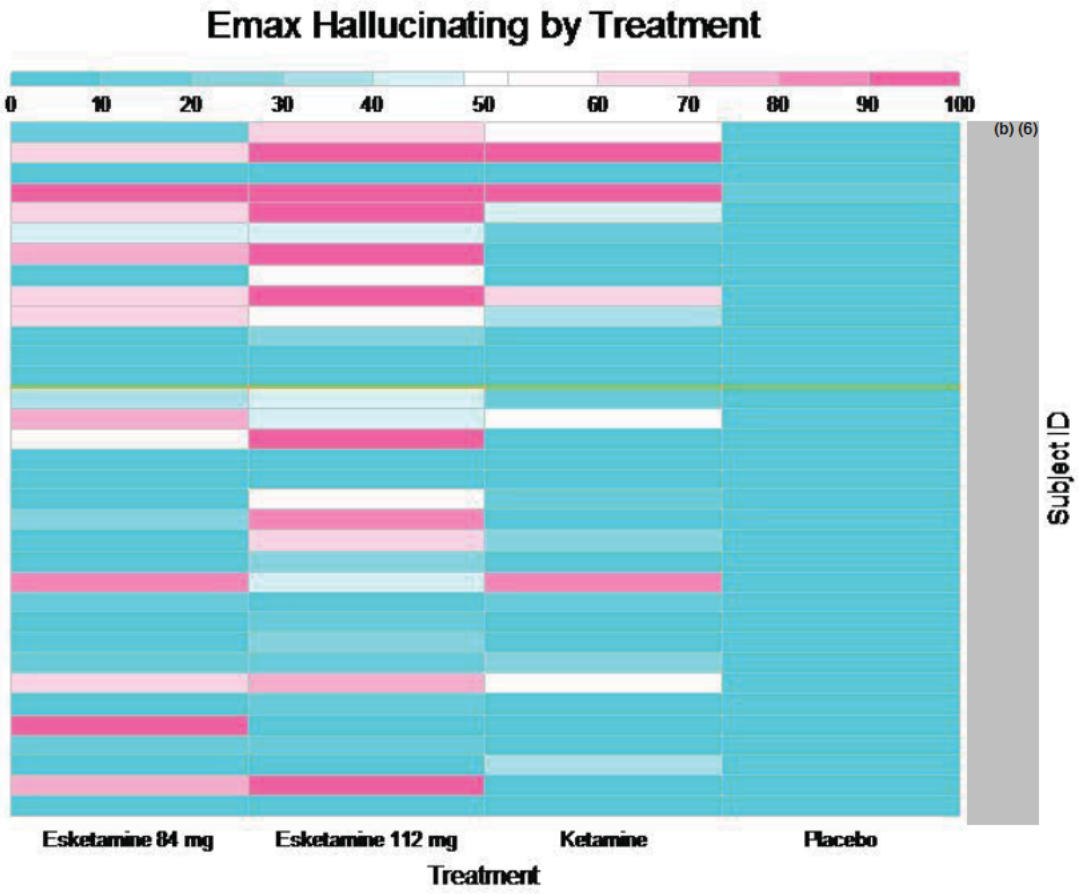


Figure 7: Heat map by treatment for Hallucinating Emax

Reference ID: 4389574

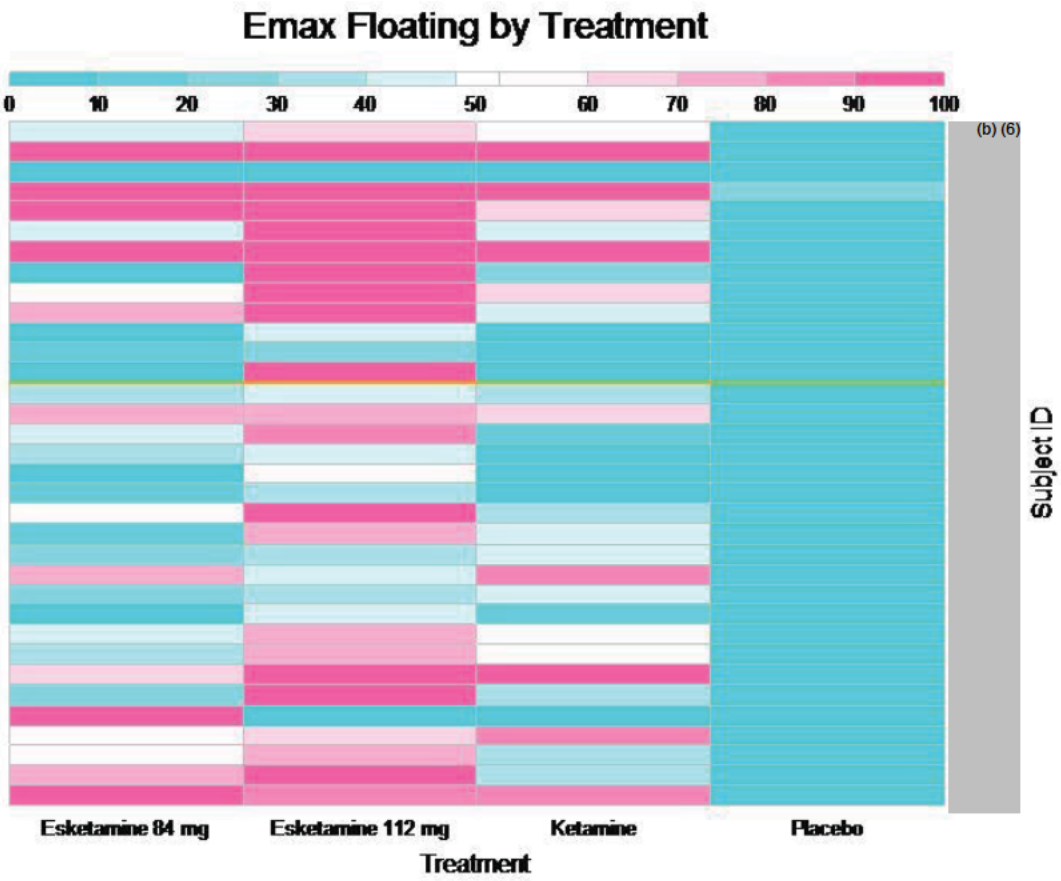


Figure 8: Heat map by treatment for Floating Emax

Reference ID: 4389574

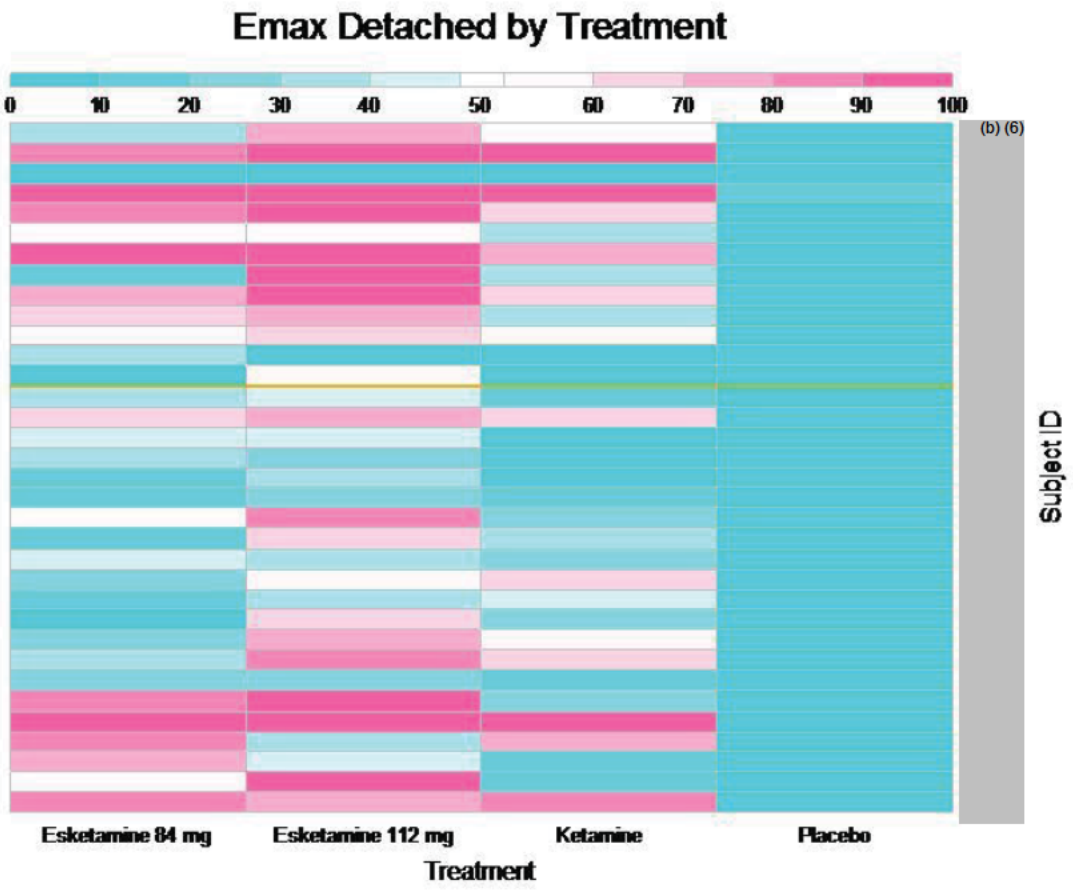


Figure 9: Heat map by treatment for Detached Emax

Reference ID: 4389574

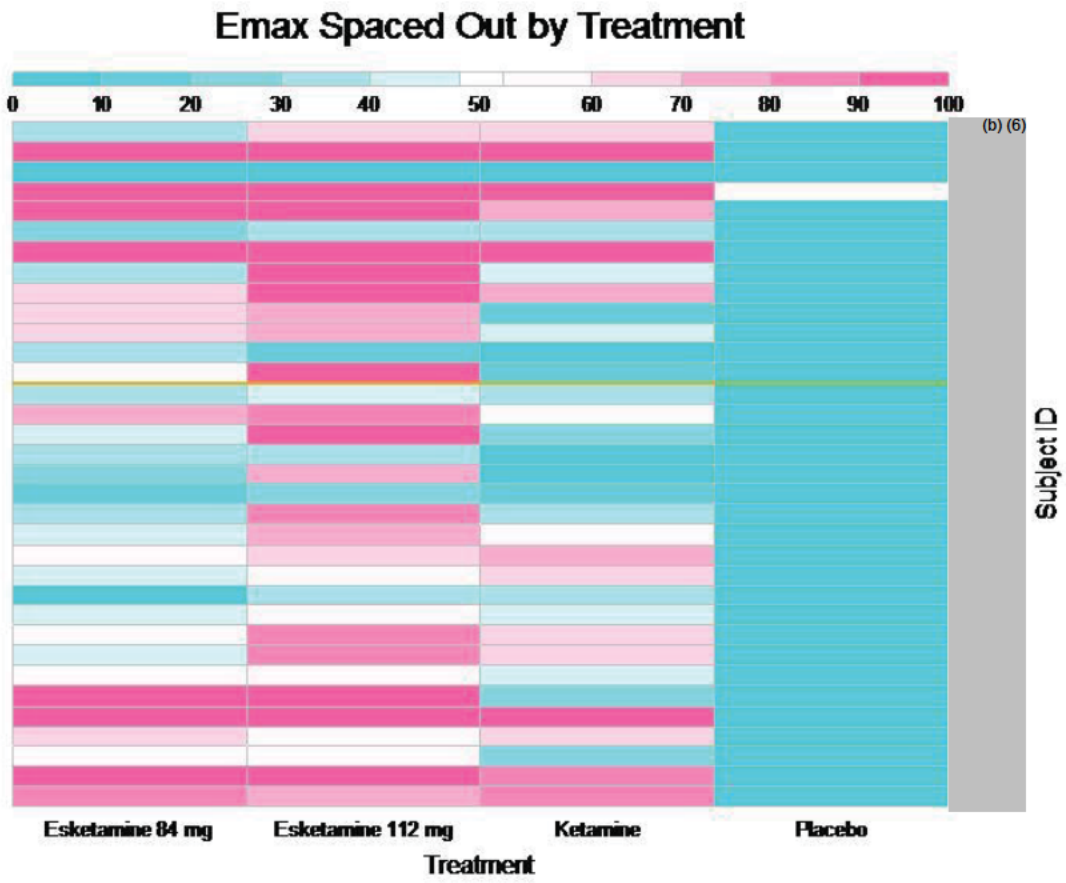


Figure 10: Heat map by treatment for Spaced Out Emax

Table 3 summarizes the mean, standard deviation (SD), minimum (Min), the first quartile (Q1), median (Med), the third quartile (Q3), and maximum (Max) for the 4 treatments in the study for the secondary endpoints of Emax Hallucinating, Floating, Detached and Spaced out.

Table 4: Summary statistics for Emax Hallucinating, Floating, Detached and Spaced out (N=34)

Measure	TRT	Mean	SD	Min	Q1	Med	Q3	Max
Hallucinating VAS N=34	Esketamine 112 mg	43.3	37.6	0	2	44	73	100
	Esketamine 84 mg	31.1	34.9	0	0	12.5	64	100
	Racemic ketamine	22.5	29.9	0	0	6.5	40	100
	Placebo	0.79	3.4	0	0	0	0	19
Floating VAS N=34	Esketamine 112 mg	69.4	30.2	0	46	73	100	100

Reference ID: 4389574

	Esketamine 84 mg	47.8	47.8	0	18	46	77	100
	Racemic ketamine	43.9	43.9	0	15	43	63	100
	Placebo	1.2	5.2	0	0	0	0	30
Detached VAS N=34	Esketamine 112 mg	63.3	30.2	0	39	67	97	100
	Esketamine 84 mg	47.7	31.2	0	22	44.5	75	100
	Racemic ketamine	41.1	31.3	0	17	36	67	100
	Placebo	0.65	2.5	0	0	0	0	14
Spaced Out VAS N=34	Esketamine 112 mg	70.9	28.4	0	52	75.5	99	100
	Esketamine 84 mg	56.5	28.1	9	35	51.5	72	100
	Racemic ketamine	49.3	30.5	0	26	50	71	100
	Placebo	1.9	9.6	0	0	0	0	56

Figure 5 and Figure 6 are plots for mean responses of Hallucinating, Floating, Detached and Spaced out, respectively. From these graphs, one may see IN esketamine 112 mg had higher mean scores in the first hours after drug dosing than IV racemic ketamine .

Reference ID: 4389574

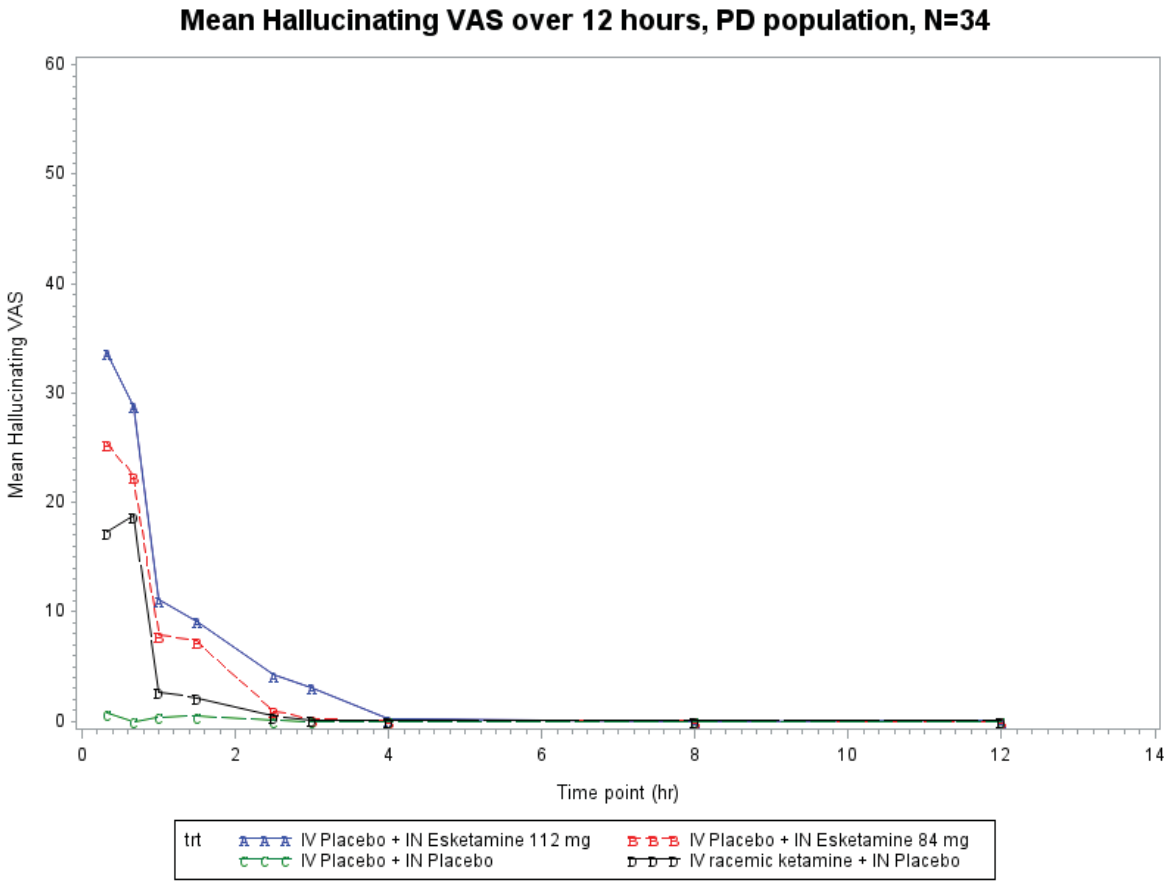


Figure 11: The mean time course profiles in 12 hours on Hallucinating VAS by treatment (N=34)

Reference ID: 4389574

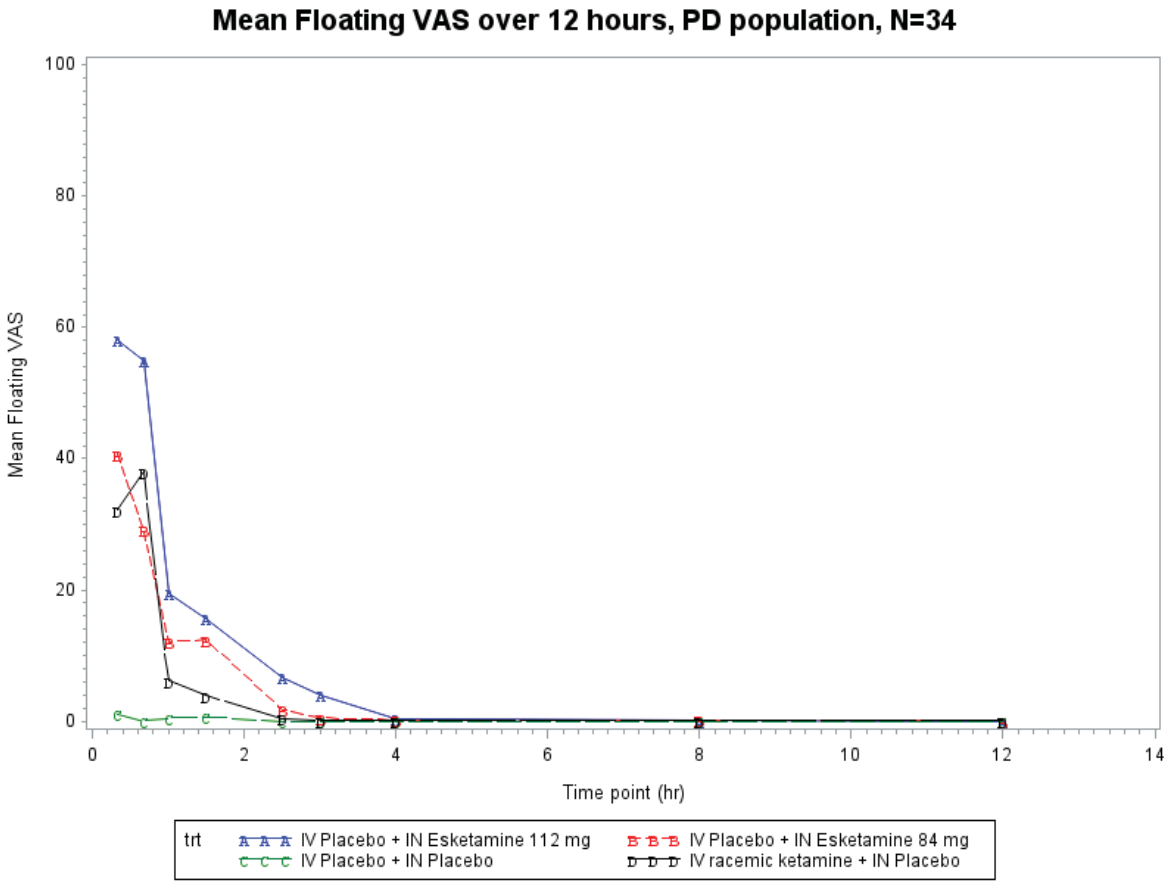


Figure 12: The mean time course profiles in 12 hours on Floating VAS by treatment (N=34)

Reference ID: 4389574

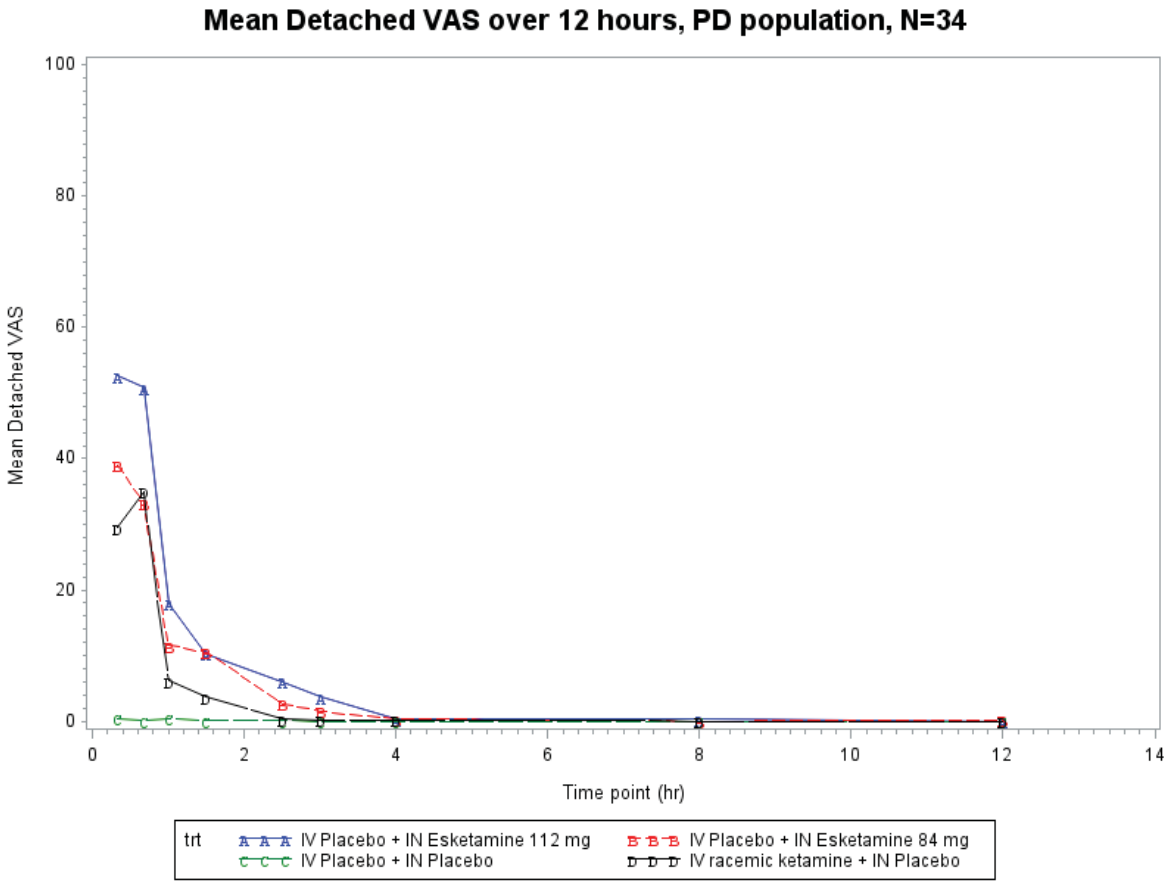


Figure 13: The mean time course profiles in 12 hours on Detached VAS by treatment (N=34)

Reference ID: 4389574

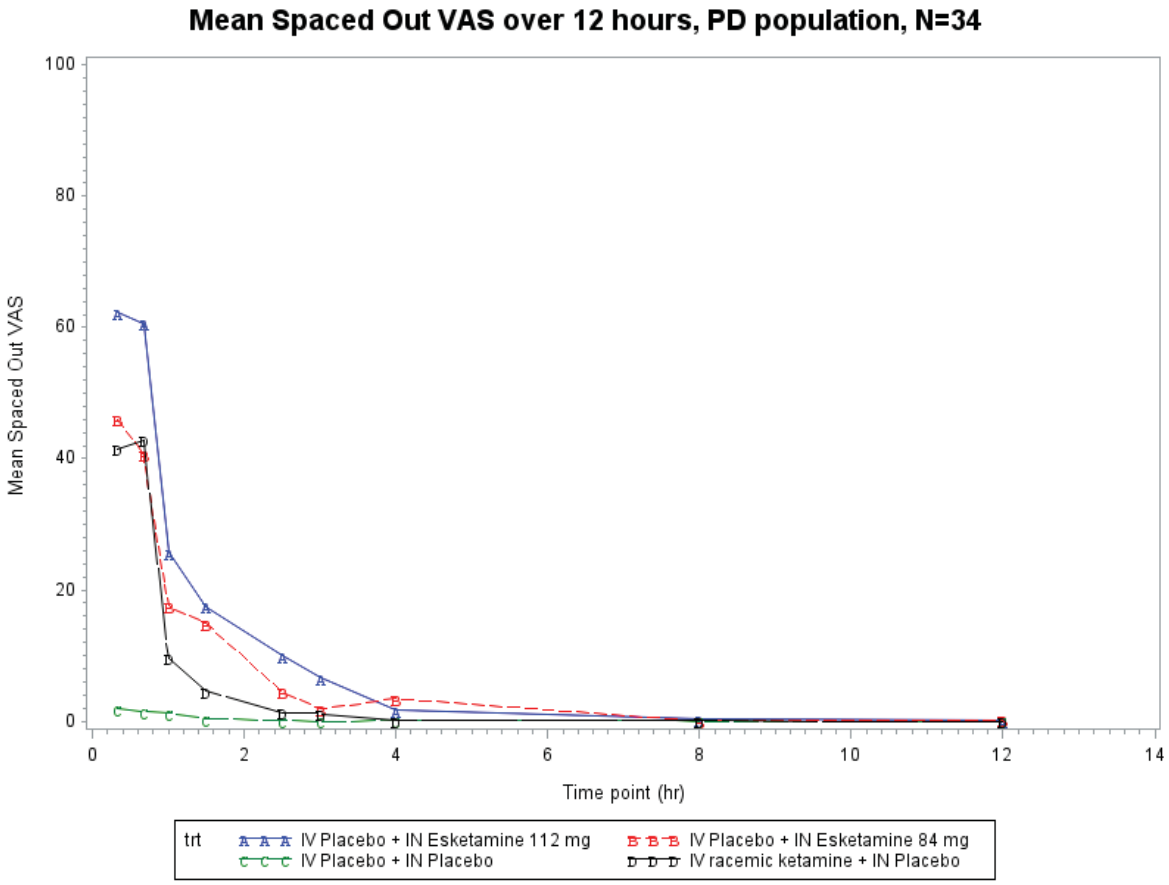


Figure 14: The mean time course profiles in 12 hours on Spaced Out VAS by treatment (N=34)

2.4.2.2. Statistical Testing of Secondary Outcomes

The statistical model used in the reviewer's secondary analysis was the mixed-effects model which included sequence, period and treatment as fixed effects, and subject as a random effect. The model was adjusted heteroscedasticity. Pre-dose responses supposed to be collected and included in the model as a covariate. However, most subjects had predose score of 0. Therefore, the predose response was not included as a covariate in the model in this analysis. For those outcomes with p-values of the Shapiro-Wilk W-test on the residuals less than 0.01, the normality of distributions of paired differences was further examined. Since all the tests performed here are either upper-tailed or two- tailed, and the paired difference distribution for those did not met the assumption of mixed model (Hallucinating, Floating, Detached and Spaced out) were either positive skewed or not very skewed (absolute value of skewness < 1), paired t tests were used. All the validation tests used margin of 0. For the comparisons between IN esketamine and IV racemic ketamine in Hallucinating, Floating, Detached and Spaced out, two-sided tests with the type I error rate of 0.1 were performed.

Reference ID: 4389574

Table 5 -- Table 10 show the results from reviewer's statistical analysis for Emax of Take Drug Again, Overall Drug Liking, Hallucinating, Floating, Detached and Spaced out, respectively.

Table 5: Statistical analysis results on Take Drug Again Emax (N=34)

Pairwise Comparison	Mean Diff /Med Diff	StdErr /IQR	p-value	95% CI	
				LCL	UCL
Racemic ketamine -P	26.4	3.4	<.001	20.8	Infty
Racemic ketamine - Esketamine 112 mg	0.325	4.5	0.47	-7.14	Infty
Racemic ketamine - Esketamine 84 mg	0.183	4.2	0.48	-6.84	Infty

Table 6: Statistical analysis results on Overall Drug Liking Emax (N=34)

Pairwise Comparison	Mean Diff /Med Diff	StdErr /IQR	p-value	95% CI	
				LCL	UCL
Racemic ketamine -P	23.8	3.2	<.001	18.5	Infty
Racemic ketamine - Esketamine 112 mg	0.85	4.5	0.43	-6.67	Infty
Racemic ketamine - Esketamine 84 mg	0.50	4.1	0.45	-6.26	Infty

Table 7: Statistical analysis results on Hallucinating Emax (N=34)

Pairwise Comparison	Mean Diff /Med Diff	StdErr /IQR	p-value	95% CI	
				LCL	UCL
Racemic ketamine -P	22.6	5.1	<.001	14.05	Infty
Racemic ketamine - Esketamine 112 mg*	-20.7	5.9	0.0013	-30.7	-10.7
Racemic ketamine -	-8.5	5.2	0.10	-17.3	0.24

Reference ID: 4389574

Esketamine 84 mg*					

*based on two-sided paired t test with alpha = 0.1

Table 8: Statistical analysis results on Floating Emax (N=34)

Pairwise Comparison	Mean Diff /Med Diff	StdErr /IQR	p-value	95% CI	
				LCL	UCL
Racemic ketamine -P	43.2	5.4	<.001	34.2	Infty
Racemic ketamine - Esketamine 112 mg*	-25.2	7.4	0.0009	--37.5	-13.0
Racemic ketamine - Esketamine 84 mg*	-3.1	7.8	0.69	-16.1	9.8

*based on two-sided test with alpha = 0.1

Table 9: Statistical analysis results on Detached Emax (N=34)

Pairwise Comparison	Mean Diff /Med Diff	StdErr /IQR	p-value	95% CI	
				LCL	UCL
Racemic ketamine -P	40.8	5.4	<.001	31.9	Infty
Racemic ketamine - Esketamine 112 mg*	-21.7	7.5	0.0045	-34.1	-9.3
Racemic ketamine - Esketamine 84 mg*	-6.3	7.6	0.41	-19.0	6.3

*based on two-sided test with alpha = 0.1

Table 10: Statistical analysis results on Spaced Out Emax (N=34)

Pairwise Comparison	Mean Diff /Med Diff	StdErr /IQR	p-value	95% CI	
				LCL	UCL
Racemic ketamine -P	47.8	5.3	<.001	39.1	Infty
Racemic ketamine - Esketamine 112 mg*	-21.0	7.0	0.0036	-32.7	-9.3
Racemic	-6.5	6.9	0.34	-17.9	4.9

Reference ID: 4389574

ketamine - Esketamine 84 mg*					

APPEARS THIS WAY ON ORIGINAL

*based on two-sided test with alpha = 0.1

The reviewer's secondary analysis showed that for the secondary outcomes:

- For all the secondary outcomes, the tests between IV racemic ketamine and placebo had $p < 0.001$, indicating study validation in all.
- there were no statistically significant differences in means between both doses of IN esketamine and IV racemic ketamine for Emax Take Drug Again and Overall Drug Liking;
- there were no statistically significant differences in means between IN esketamine 84 mg and IV racemic ketamine for Emax Hallucinating, Floating, Detached and Spaced out;
- IN esketamine had significantly higher score in Emax Hallucinating, Floating, Detached and Spaced out than IV racemic ketamine.

3. Conclusion

In conclusion, the measure of Emax Drug Liking, Take Drug again and Overall Drug liking after IN esketamine 84 mg and 112 mg dosing are comparable with IV racemic ketamine. However, IN esketamine 112 mg had significantly higher maximum VAS score than IV racemic ketamine in Hallucinating, Floating, Detached and Spaced out, indicating possible higher perceptual effects with high dose of esketamine.

Reference ID: 4389574

This is a representation of an electronic record that was signed electronically. Following this are manifestations of any and all electronic signatures for this electronic record.

/s/

QIANYU DANG
02/12/2019 11:55:40 AM

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Science
Office of Biostatistics

Statistical Review and Evaluation

CARCINOGENICITY STUDIES

IND/NDA Number:	NDA 211243
Drug Name:	JNJ-54135419-AAC (hydrochloride HCl salt)
Indication:	Treatment-resistant depression
Studies:	Carcinogenicity Studies in Rats for 104 Weeks and Mice for 26 Weeks
Applicant:	Sponsor: Janssen Research & Development Turnhoutseweg 30 B-2340 Beerse, Belgium Testing Facility for Rats: (b) (4) Testing Facility for Mice: (b) (4)
Review Priority:	Standard
Biometrics Division:	Division of Biometrics - VI
Statistical Reviewer:	Hepei Chen
Concurring Reviewer:	Karl Lin, Ph.D.
Medical Division:	Division of Psychiatry Products
Reviewing Pharmacologist:	Shiny Mathew, Ph.D.
Keywords:	Carcinogenicity, Dose response

Table of Contents

1. Background 3
2. Rat Study 3
   2.1. Sponsor's analyses 4
      2.1.1. Survival analysis 4
         Sponsor's findings 4
      2.1.2. Tumor data analysis 4
         Adjustment for multiple testing 4
         Sponsor's findings 4
   2.2. Reviewer's analyses 5
      2.2.1. Survival analysis 5
         Reviewer's findings 6
      2.2.2. Tumor data analysis 6
         Multiple testing adjustment 6
         Reviewer's findings 7
3. Mouse Study 8
   3.1. Sponsor's analyses 8
      3.1.1. Survival analysis 9
         Sponsor's findings 9
      3.1.2. Tumor data analysis 9
         Sponsor's findings 10
   3.2. Reviewer's analyses 10
      3.2.1. Survival analysis 11
         Reviewer's findings 11
      3.2.2. Tumor data analysis 11
         Reviewer's findings 11
4. Summary 11
5. Appendix 14
   Table 1A: Intercurrent mortality rate in male rats
   Table 1B: Intercurrent mortality rate in female rats
   Table 2A: Tumor rates and p-values for trend and pairwise comparisons in male rats
   Table 2B: Tumor rates and p-values for trend and pairwise comparisons in female rats
   Table 3A: Intercurrent mortality rate in male mice
   Table 3B: Intercurrent mortality rate in female mice
   Table 4A: Tumor rates and p-values for trend and pairwise comparisons in male mice
   Table 4B: Tumor rates and p-values for trend and pairwise comparisons in female mice
   Figure 1A: Kaplan-Meier survival functions for male rats
   Figure 1B: Kaplan-Meier survival functions for female rats
   Figure 2A: Kaplan-Meier survival functions for male mice
   Figure 2B: Kaplan-Meier survival functions for female mice
6. References 32

1. Background

In this submission the sponsor included reports of two animal carcinogenicity studies, one in rats and one in mice. These studies were to evaluate the oncogenic potential of the test article, JNJ-54135419-AAC (hydrochloride HCl salt), an NMDA receptor antagonist, when administered daily via intranasal instillation in rats for 104 weeks, and via subcutaneous administration in mice for 26 weeks.

In this review the phrase "dose response relationship" refers to the linear component (trend) of the effect of treatment, and not necessarily to a strictly increasing or decreasing mortality or tumor incidence rate as dose increases.

2. Rat Study

Two separate experiments, one in male rats and one in female rats were conducted. As indicated in Table 1, in each of these two experiments there were three treated groups, one saline control group, and one vehicle control group. Three hundred twenty five Sprague-Dawley rats of each sex were assigned randomly in size of 65 rats per group. The dose levels for the three treated groups were 0.9, 3, and 9 mg/kg/day for male and female rats. In this review these dose groups were referred to as the low (Group 3), mid (Group 4), and high (Group 5) dose groups, respectively. The rats in the saline and vehicle control groups were administrated with physiological saline and vehicle [water for irrigation, citric acid, EDTA and pH 4.5 (via NaOH)], respectively, and handled for the same duration and in the same manner as the treated groups.

Table 1: Experimental Design in Rat Study

Group No.	No. of Animals		Test Material	Dosage Level (mg/kg/day)	
	Male	Female		Male	Female
1	65	65	Saline Control	0	0
2	65	65	Vehicle control	0	0
3	65	65	JNJ-54135419-AAC Low	0.9	0.9
4	65	65	JNJ-54135419-AAC Mid	3	3
5	65	65	JNJ-54135419-AAC High	9	9

With only 20 main study males surviving for group 1 (saline control) on 16 Nov 2015 (Week 102), male rats of all groups were terminated at the earliest opportunity (Week 102/103). Female rats (all groups) were terminated in Week 105.

All animals were observed once during the pretreatment period, then at least once daily during the dosing period, for ill health, abnormal behaviour or unusual appearance, untoward clinical signs, toxic or pharmacological response. The size, appearance, position and duration of any masses detected was recorded in Week -1, on Day 1 and weekly thereafter. When possible, the animals were euthanised rotating across dose groups such that similar numbers of animals from each group, including controls were necropsied throughout the day. Animals were not fasted before their scheduled necropsy. All main study animals were subjected to a complete necropsy examination, which included evaluation of the carcass and musculoskeletal system; all external surfaces and orifices; cranial cavity and external surfaces of the brain; and thoracic, abdominal, and pelvic cavities with their associated organs and tissues.

2.1. Sponsor's analyses

2.1.1. Survival analysis

In the sponsor's analysis, Kaplan-Meier estimated survival functions were calculated for the saline control, vehicle control and each of 3 dose groups of the test material. Intergroup comparisons of mortality comparing the saline control and each treatment group with the vehicle control group were performed using a Wilcoxon rank sums test modified for censored observations..

Sponsor's findings:

The sponsor's analysis showed that the numbers of rats surviving to their terminal necropsy were 20 (31%), 22 (34%), 23 (35%), 25 (38%), and 23 (35%) in Groups 1, 2, 3, 4, and 5 for male rats, respectively, and 21 (32%), 21 (32%), 20 (31%), 16 (25%), and 21 (32%) for female rats respectively. The sponsor's report indicated that there were no test article-related difference in mortality between control and treated male and female rats.

2.1.2. Tumor data analysis

In the sponsor's analysis, the study pathologist classified each of the tumors as fatal, probably fatal, probably incidental or incidental for all tumors. For the purposes of statistical analysis of non-palpable tumors, tumors classified as either fatal or probably fatal were considered fatal and those classified as probably incidental or incidental were considered incidental. All tumors detected in animals that died in a planned sacrifice were automatically classified as incidental.

For the purposes of statistical analysis, all observable tumors (i.e. mammary gland and skin tumors) were considered to be fatal and the time of first detection was used in the age-adjusted analysis. For observable tumors not detected in life, the time to death was used in the analysis. The analysis of incidental tumors was conducted by dividing the experimental period into the following fixed time intervals: 1-52 weeks, 53-78 weeks, 79-92 weeks, over 92 weeks and a single interval for the planned sacrifice.

For each considered dataset, the significance of a linear dose related increase in tumor incidence was evaluated using a one-sided trend test. For the purposes of the trend test, Groups 2 to 5 only were considered and the dose levels were used. Furthermore, Pete's one-sided test was used to test whether or not the tumor incidence in each test item treated group was significantly higher than the vehicle control. An additional two-sided comparison of tumor incidence in Groups 1 and 2 was also performed.

For each statistical test performed on a dataset containing 10 or less tumors, the discrete permutation distribution was used to calculate the corresponding p-value.

Adjustment for multiple testing:

No adjustment for multiple testing was provided in the sponsor's report.

Sponsor's findings:

Significant results (at 5% significance level) have been summarised below:

Organ	Tumour Type	Sex	P(1)	P(2)	P(3)	P(4)	P(5)
skin	lipoma	M	0.31	0.93	0.044	0.058	0.85
		F	1.00	0.94	0.77	1.00	1.00
pancreas	Islet cell adenoma	M	0.50	0.56	0.009	0.017	0.076
		F	0.67	0.42	0.94	1.00	0.76
gland, thyroid	c-cell adenoma	M	0.69	0.57	0.18	0.071	0.37
		F	0.40	0.49	0.18	0.014	0.27
Uterus	Endometrial stromal polyp	F	0.36	0.008	0.90	0.16	0.046

(1) P-value for pairwise comparison of saline control with Vehicle Control
(2) P-value for linear trend including groups 2 to 5
(3) P-value for pairwise comparison of 0.9 mg eq./day JNJ-54135419-AAC with Vehicle Control
(4) P-value for pairwise comparison of 3 mg eq./day JNJ-54135419-AAC with Vehicle Control
(5) P-value for pairwise comparison of 9 mg eq./day JNJ-54135419-AAC with Vehicle Control

The sponsor's analysis showed that the incidence of skin lipoma in male rats was significantly higher in the low dose group compared with the vehicle control (p=0.044). However, the test for a dose related increase was not significant (p=0.93). There was no significant increase in this finding for female rats.

The incidence of pancreas islet cell adenoma in male rats was significantly higher in the low and mid dose groups compared with the vehicle control (p=0.009 and 0 .017 respectively). However, the test for a dose related increase was not significant (p=0.56). There was no significant increase in this finding for female rats.

The incidence of thyroid gland c-cell adenoma in female rats was significantly higher in the mid dose group compared with the vehicle control (p=0.014). However, the test for a dose related increase was not significant (p=0.49). There was no significant increase in this finding for male rats.

The incidence of uterus endometrial stromal polyp was significantly higher in the high dose group compared with the vehicle control (p=0.046). In addition, the test for a dose related increase was significant (p=0.008).

No other statistically significant findings were noted for both male and female rats.

2.2. Reviewer's analyses

To verify the sponsor's analyses and to perform additional analyses suggested by the reviewing toxicologist, this reviewer independently performed the survival and tumor data analyses using the data provided by the sponsor electronically.

2.2.1. Survival analysis

In the reviewer's analysis, the survival distributions of rats in all five groups (Groups 1, 2, 3, 4, and 5) were estimated using the Kaplan-Meier product limit method. The dose response relationship was tested across Groups 2, 3, 4, and 5 using the likelihood ratio test, and the homogeneity of survival distributions was tested using the log-rank test. The Kaplan-Meier curves for survival rates are given in Figures 1A and 1B in the appendix for all five groups in male and female rats, respectively. The intercurrent mortality data of all five groups and the results of the tests for dose response relationship and homogeneity of survivals for Groups 2, 3, 4, and 5 are given in Tables 1A and 1B

in the appendix for male and female rats, respectively.

Reviewer's findings:

The reviewer's analysis showed that the numbers of rats surviving to their terminal necropsy were 20 (31%), 22 (34%), 23 (35%), 25 (38%), and 23 (35%) in Groups 1, 2, 3, 4, and 5 for male rats, respectively, and 21 (32%), 21 (32%), 20 (31%), 16 (25%), and 21 (32%) for female rats respectively. No statistically significant dose response relationship and pairwise comparisons in mortality was noted for both male and female rats.

2.2.2. Tumor data analysis

The tumor data were analyzed for dose response relationships across Groups 2, 3, 4, and 5, and pairwise comparisons of each of the three treated groups (Groups 3, 4, and 5) against the vehicle control group (Group 2), using the Poly-k method described in the paper of Bailer and Portier (1988) and Bieler and Williams (1993).

In the ploy-k method, the adjustment for differences in mortality among treatment groups is made by modifying the number of animals at risk in the denominators in the calculations of overall tumor rates in the Cochran-Armitage test to reflect less-than-whole-animal contributions for animals that die without tumor before the end of the study (Bailer and Portier 1988). The modification is made by defining a new number of animals at risk for each treatment group. The number of animals at risk for the *i*-th treatment group R^* i is defined as $R^*_i = \sum W_{ij}$ where w_{ij} is the weight for the *j*-th animal in the *i*-th treatment group, and the sum is over all animals in the group.

Bailer and Portier (1988) proposed the weight w_{ij} as follows:

$w_{ij} = 1$ to animals dying with the tumor, and

$w_{ij} = (t_{ij} / tsacr)^3$ to animals dying without the tumor,

where t_{ij} is the time of death of the *j*-th animal in the *i*-th treatment group, and tsacr is the planned (or intended) time of terminal sacrifice. The above formulas imply that animals living up to the end of the planned terminal sacrifice date without developing any tumor will also be assigned $w_{ij} = 1$ since t_{ij} = tsacr. Also animals developed the tumor type being tested before the end of the study will be assigned as $w_{ij} = 1$.

Certain treatment groups of a study or the entire study may be terminated earlier than the planned (or intended) time of terminal sacrifice due to excessive mortalities. However, based on the principle of the Intention-to-treat (ITT) analysis in randomized trials, the tsacr should not be affected by the unplanned early terminations. The tsacr should always be equal to the planned (or intended) time of terminal sacrifice. For those animals that were sacrificed later than tsacr, regardless their actual terminal sacrifice time, tsacr was used as their time of terminal sacrifice in the analysis.

One critical point for Poly-k test is the choice of the appropriate value of k, which depends on the tumor incidence pattern with the increased dose. For long term 104 week standard rat and mouse studies, a value of k=3 is suggested in the literature. Hence, this reviewer used k=3 for the analysis of this data.

Multiple testing adjustment:

For the adjustment of multiple testing, this reviewer used the methodologies suggested in the FDA guidance for statistical design and analysis of carcinogenicity studies (2001). For dose response relationship tests, the guidance suggests the use of test levels of α=0.01 for common tumors and α=0.05 for rare tumors for a submission with one two-year study in one species and one short-term study with another species, in order to keep the overall false-positive rate at the nominal level of approximately 10%. For multiple pairwise comparisons of treated group with control group, however, the guidance indicated that the corresponding multiple testing adjustment is still under development and not yet available. To be conservative, the test level of α=0.05 was used for pairwise comparisons of treated group with control group for both rare and common tumors in this study.

It should be noted that the FDA guidance for multiple testing for dose response relationship is based on a publication by Lin and Rahman (1998). In this work the authors investigated the use of this rule for Peto analysis. However, in a later work Rahman and Lin (2008) showed that this rule for multiple testing for dose response relationship is also suitable for Poly-k tests.

A rare tumor is defined as one in which the published spontaneous tumor rate is less than 1%. However, if the background information for the common or rare tumor is not available, the number of animals bearing tumors in the vehicle control group in the present study was used to determine the common or rare tumor status in the review report. That is, if the number of animals bearing tumors in the vehicle control group is 0, then this tumor is considered as the rare tumor; otherwise, if the number of animals bearing tumors in the control group is greater than or equal to 1, then this tumor is considered as the common tumor.

Reviewer's findings:

The tumor rates and the p-values of the tested tumor types are listed in Tables 2A and 2B in the appendix for male and female rats, respectively. The tumor types with p-values less than or equal to 0.05 for dose response relationship and/or pairwise comparisons of treated groups and vehicle control are reported in Table 2.

Based on the criteria of adjustment for multiple testing discussed above, the reviewer's analysis showed statistically significant positive dose response relationship for the incidence of interstitial (leydig) cell adenoma in testis of male rats (p-value =0.0491) if this tumor is considered to be rare, without any corresponding statistically significant pairwise comparison between the treated groups and the vehicle control group. Statistically significant increases for the incidence of islet cell adenoma in pancreas were noted in the low and mid dose groups in male rats when comparing to the vehicle control group (p-value = 0.0312 and 0.0376, respectively) regardless the classification of this tumor (rare or common).

For female rats, statistically significant increases were noted in the mid dose group for the incidence of combined adenoma and carcinoma in pituitary gland (p-value = 0.0494) and the incidence of combined c-cell adenoma and carcinoma in thyroid gland (p-value = 0.0223) when comparing to the vehicle control group, regardless the classification of this tumor (rare or common). No corresponding statistically significant dose response relationship was noted for these tumors.

No other statistically significant findings were noted in tumor data for both male and female rats.

Table 2. Summary Table of Tumor Types with P-Values ≤ 0.05 for Dose Response Relationship and/or Pairwise Comparisons of Treated Groups and Vehicle control Group in Rats

Organ name	Tumor name	Vehicle (VC) 0 mg P - Trend	Low (L) 0.9 mg P - VC vs. L	Mid (M) 3 mg P - VC vs. M	High (H) 9 mg P - VC vs. H	Saline (SC) 0 mg P - VC vs. SC
Male						
Pancreas	Islet Cell Adenoma	0/62 (38)	5/62 (40)	5/64 (43)	2/64 (40)	1/63 (39)
		0.5520	0.0312 $	0.0376 $	0.2597	0.5065
Testis	Interstitial (Leydig) Cell Adenoma	0/65 (40)	1/65 (40)	1/65 (43)	3/65 (43)	1/65 (40)
		0.0491 $	0.5000	0.5181	0.1343	0.5000
Female						
Gland, Pituitary	Adenoma	43/65 (59)	50/64 (62)	51/65 (59)	41/63 (56)	49/64 (59)
		0.6926	0.2129	0.0541	0.5677	0.1333
	Carcinoma	0/65 (43)	0/64 (43)	1/65 (40)	0/63 (40)	1/64 (43)
		0.4819	NC	0.4819	NC	0.5000
	Adenoma/Carcinoma	43/65 (59)	50/64 (62)	52/65 (60)	41/63 (56)	50/64 (59)
		0.6861	0.2129	0.0494 $	0.5677	0.9121
Gland, Thyroid	C-Cell Adenoma	1/65 (43)	3/63 (43)	7/63 (40)	2/65 (41)	3/64 (44)
		0.4777	0.3080	0.0223 $	0.4819	0.3168
	C-Cell Carcinoma	0/65 (43)	0/63 (43)	0/63 (39)	1/65 (41)	0/64 (43)
		0.2470	NC	NC	0.4881	NC
	C-Cell Adenoma/ C-Cell Carcinoma	1/65 (43)	3/63 (43)	7/63 (40)	3/65 (41)	3/64 (44)
		0.3075	0.3080	0.0223 $	0.2900	0.6832
Uterus	Endometrial Stromal Polyp	2/65 (44)	0/64 (43)	4/64 (39)	6/63 (40)	4/65 (44)
		0.0103 @	1.0000	0.2819	0.1039	0.3382

& X/ZZ (YY): X=number of tumor bearing animals; YY=mortality weighted total number of animals; ZZ=unweighted total number of animals observed;
$ = Statistically significant at 0.05 level for test of dose response relationship in rare tumor and for test of pairwise comparisons regardless tumor classification (rare or common);
@ = Not statistically significant at 0.01 level for test of dose response relationship in common tumor;
NC = Not calculable.

3. Mouse Study

Table 3. Experimental Design in Mouse Study

Group No.	No. of Animals		Test Material	Dosage Level (mg/kg/day)	
	Male	Female		Male	Female
1	25	25	Saline Control	0	0
2	25	25	Vehicle control	0	0
3	25	25	JNJ-54135419-AAC Low	10	10
4	25	25	JNJ-54135419-AAC Mid	25	25
5	25	25	JNJ-54135419-AAC High	75/40*	75/40*
6	10	10	Positive Control	0	0

* Due to observed toxicity, the dose level was decreased to 40 mg/kg/day beginning on Day 124 for male mice or Day 122 for female mice.

Two separate experiments, one in male mice and one in female mice were conducted. As indicated in Table 3, in each of these two experiments there were three treated groups, one saline control group, one positive control group, and one vehicle control group. One hundred and thirty

five Tg.rasH2 mice of each sex were assigned randomly in size of 25 mice per group except for the positive control group with 10 mice of each sex. The dose levels for the three treated groups were 10, 25, and 75(40) mg/kg/day for both male and female mice. In this review these dose groups were referred to as the low (Group 3), mid (Group 4), and high (Group 5) dose groups, respectively. The mice in the saline control and vehicle control group were administrated with normal sterile water, and vehicle [1.5 mg/mL citric acid and 0.12 mg/mL EDTA, in sterile water for injection, pH 4.5 ± 0.1 (adjusted with 1N NaOH)], and handled for the same duration and in the same manner as the treated groups.

Observations for morbidity, mortality, injury, and the availability of food and water were conducted twice daily for all animals. Observations for clinical signs and masses were conducted once weekly on main study animals. Body weights for all animals were measured and recorded pretest and weekly. Food consumption was measured and recorded weekly for main study animals. Ophthalmoscopic examinations were conducted pretest and prior to the terminal necropsy for main study animals. Blood samples for determination of the plasma concentrations of the test article were collected from TK animals at designated time points during Month 3, Month 6, and on Day 368. After blood collection, the TK animals were euthanized and the carcasses were discarded without further evaluation. At study termination, necropsy examinations were performed and selected tissues were microscopically examined.

3.1. Sponsor's analyses

3.1.1. Survival analysis

In the sponsor’s report, Kaplan-Meier estimates of group survival rates were calculated, by sex, and shown graphically. For each control group (saline and vehicle control), the generalized Wilcoxon test for survival was used to compare the homogeneity of survival rates across the control and test article groups, by sex, at the 0.05 significance level. If the survival rates were significantly different, the generalized Wilcoxon test was used to make pairwise comparisons of each test article group with each of the control groups. Additionally, the positive control group was compared to the saline and vehicle control groups using the generalized Wilcoxon test.

Survival times in which the status of the animal’s death was classified as an accidental death, planned interim sacrifice or terminal sacrifice were considered censored values for the purpose of the Kaplan-Meier estimates and survival rate analyses.

Sponsor’s findings:

The sponsor’s analysis showed that the numbers of mice surviving to their terminal necropsy were 25 (100%), 25 (100%), 25 (100%), 22 (88%), and 18 (72%) in Groups1, 2, 3, 4, and 5 for male mice, respectively, and 24 (96%), 24 (96%), 22 (88%), 25 (100%), and 21 (84%) for female mice, respectively. Among males, there was a statistically significant increase in mortality rates when comparing the high dose group with both saline and vehicle control groups. There were no other statistically significant findings in survival rates.

3.1.2. Tumor data analysis

In the sponsor’s analysis, the incidence of tumors was analyzed by Peto’s mortality-prevalence method, without continuity correction, incorporating the context (incidental, fatal, or mortality independent) in which tumors were observed. Because of the sparse number of deaths during the

study, the following fixed intervals were used for incidental tumor analyses: Days 1 through 100 and Days 101 through and including terminal sacrifice. A minimum exposure of 101 days was considered sufficient to be included with animals surviving through scheduled termination. All tumors in the scheduled terminal sacrifice interval were considered incidental for the purpose of statistical analysis.

Tumors classified as mortality-independent were analyzed with Peto's mortality independent method incorporating the day of detection. Each diagnosed tumor type was analyzed separately and, at the discretion of the study director, analysis of combined tumor types and/or organs was performed. All metastases and invasive tumors were considered secondary and not included in the analyses.

A 1-sided comparison of each test article group with the saline and vehicle control was performed. The saline and vehicle control groups were compared with a 2-sided test. An exact permutation test was conducted for all analyses. Findings were evaluated for statistical significance at both the 0.01 and 0.05 levels.

Because the positive control group was scheduled for early terminal sacrifice, tumor incidence in the positive control group was compared to the saline and vehicle control groups with a 1-sided Fisher's exact test at both the 0.01 and 0.05 significance levels and all p values were reported. Only the following tumors were statistically analyzed: alveolar-bronchiolar adenoma, alveolar-bronchiolar carcinoma, and hemangiosarcoma in the spleen.

Multiple testing adjustment:

No adjustment for multiple testing was provided in the sponsor's report.

Sponsor's findings:

In the sponsor's analysis, no statistically significant tumor findings in the test article groups were noted when compared to the saline and vehicle control groups for both male and female mice. There was a statistically significant increase in the following tumors when comparing the positive control with both the saline and vehicle control group:

Sex	Organ	Tumor
M	LUNGS WITH BRONCHI	ALVEOLAR-BRONCHIOLAR ADENOMA (B)
		CARCINOMA/ADENOMA
F	LUNGS WITH BRONCHI	ALVEOLAR-BRONCHIOLAR ADENOMA (B)
		CARCINOMA/ADENOMA
(B)- Benign; (M)- Malignant		

.

3.2. Reviewer's analyses

Similar to the rat study, this reviewer independently performed survival and tumor data analyses of mouse data to verify sponsor's analyses. Data used in this reviewer's analyses were provided by the sponsor electronically.

In all statistical analysis, because the dose level of the high dose group was decreased to 40 mg/kg/day beginning on Day 124 for male mice or Day 122 for female mice, the dose level score was adjusted by taking the weighted arithmetic mean of doses during the different dosing periods. The corresponding adjusted dose level score for the high dose group used in the statistical analysis was 64 mg/kg/day for both male and female mice.

For the analysis of both the survival data and the tumor data in mice, this reviewer used similar methodologies that were used for the analyses of the rat survival and tumor data.

3.2.1. Survival analysis

The Kaplan-Meier curves for survival rates of all treatment groups are given in Figures 2A and 2B in the appendix for male and female mice, respectively. The intercurrent mortality data, and the results of the tests for dose response relationship and homogeneity of survivals for the combined vehicle control, low, mid, and high dose groups were given in Tables 3A and 3B in the appendix for male and female mice, respectively.

Reviewer's findings:

The reviewer's analysis showed that the numbers of mice surviving to their terminal necropsy were 25 (100%), 25 (100%), 25 (100%), 22 (88%), and 18 (72%) in Groups 1, 2, 3, 4, and 5 for male mice, respectively, and 24 (96%), 24 (96%), 22 (88%), 25 (100%), and 21 (84%) for female mice, respectively. The reviewer's analysis also showed a statistically significant dose response relationship in mortality in male mice (p-value = 0.0003), with the corresponding statistically significant increases in the mid and high dose groups in mortality when comparing to the vehicle control group (p-value = 0.0384 and 0.0011, respectively). No significant findings were noted in mortality for female mice.

3.2.2. Tumor data analysis

The tumor rates and the p-values of the tested tumor types are given in Tables 4A and Table 4B in the appendix, for male and female mice, respectively.

Reviewer's findings:

The reviewer's analysis showed no statistically significant dose response relationship or pairwise comparisons in the treated groups when compared to the vehicle control group for both male and female mice. When comparing the positive control with the vehicle control group, statistically significant increases in the incidence of alveolar-bronchiolar adenoma, and the combined alveolar-bronchiolar adenoma and carcinoma in lungs with bronchi were noted for both male and female mice (all p-values < 0.0001).

4. Summary

In this submission the sponsor included reports of two animal carcinogenicity studies, one in rats and one in mice. These studies were to evaluate the oncogenic potential of the test article, JNJ-54135419-AAC (hydrochloride HCl salt), an NMDA receptor antagonist, when administered daily via intranasal instillation in rats for 104 weeks, and via subcutaneous administration in mice for 26 weeks.

Rat Study:

Two separate experiments, one in male rats and one in female rats were conducted. In each of these two experiments there were three treated groups, one saline control group, and one vehicle control group. Three hundred twenty five Sprague-Dawley rats of each sex were assigned randomly in size of 65 rats per group. The dose levels for the three treated groups were 0.9, 3, and 9 mg/kg/day for male and female rats.

The reviewer's analysis showed that the numbers of rats surviving to their terminal necropsy were 20 (31%), 22 (34%), 23 (35%), 25 (38%), and 23 (35%) in Groups 1, 2, 3, 4, and 5 for male rats, respectively, and 21 (32%), 21 (32%), 20 (31%), 16 (25%), and 21 (32%) for female rats respectively. No statistically significant dose response relationship and pairwise comparisons in mortality was noted for both male and female rats.

The reviewer's analysis showed statistically significant positive dose response relationship for the incidence of interstitial (leydig) cell adenoma in testis of male rats (p-value =0.0491) if this tumor is considered to be rare, without any corresponding statistically significant pairwise comparison between the treated groups and the vehicle control group. Statistically significant increases for the incidence of islet cell adenoma in pancreas were noted in the low and mid dose groups in male rats when comparing to the vehicle control group (p-value = 0.0312 and 0.0376, respectively) regardless the classification of this tumor (rare or common). For female rats, statistically significant increases were noted in the mid dose group for the incidence of combined adenoma and carcinoma in pituitary gland (p-value = 0.0494) and the incidence of combined c-cell adenoma and carcinoma in thyroid gland (p-value = 0.0223) when comparing to the vehicle control group, regardless the classification of this tumor (rare or common). No corresponding statistically significant dose response relationship was noted for these tumors. No other statistically significant findings were noted in tumor data for both male and female rats.

Mouse Study:

Two separate experiments, one in male mice and one in female mice were conducted. In each of these two experiments there were three treated groups, one saline control group, one positive control group, and one vehicle control group. One hundred and thirty five Tg.rasH2 mice of each sex were assigned randomly in size of 25 mice per group except for the positive control group with 10 mice of each sex. The dose levels for the three treated groups were 10, 25, and 75(40) mg/kg/day for both male and female mice.

The reviewer's analysis showed that the numbers of mice surviving to their terminal necropsy were 25 (100%), 25 (100%), 25 (100%), 22 (88%), and 18 (72%) in Groups 1, 2, 3, 4, and 5 for male mice, respectively, and 24 (96%), 24 (96%), 22 (88%), 25 (100%), and 21 (84%) for female mice, respectively. The reviewer's analysis also showed a statistically significant dose response relationship in mortality in male mice (p-value = 0.0003), with the corresponding statistically significant increases in the mid and high dose groups in mortality when comparing to the vehicle control group (p-value = 0.0384 and 0.0011, respectively). No significant findings were noted in mortality for female mice.

The reviewer's analysis showed no statistically significant dose response relationship or pairwise comparisons in the treated groups when compared to the vehicle control group for both male and

female mice. When comparing the positive control with the vehicle control group, statistically significant increases in the incidence of alveolar-bronchiolar adenoma, and the combined alveolar-bronchiolar adenoma and carcinoma in lungs with bronchi were noted for both male and female mice (all p-values < 0.0001).

Hepei Chen.
Mathematical Statistician

Concur: Karl Lin, Ph.D.
Team Leader, DBVI

Cc: Archival NDA 211243

Dr. Shiny Mathew
Dr. Lillian Patrician

5. Appendix

Table 1A: Intercurrent Mortality Rate in Male Rats

	Vehicle control		Low		Mid		High		Saline Control	
Week / Type of Death	No. of Death	Cum %	No. of Death	Cum %	No. of Death	Cum %	No. of Death	Cum %	No. of Death	Cum %
0 - 52	2	3.08	3	4.62	1	1.54	5	7.69	5	7.69
53 - 78	17	29.23	19	33.85	16	26.15	10	23.08	16	32.31
79 - 91	18	56.92	11	50.77	14	47.69	18	50.77	12	50.77
92 - 104	6	66.15	9	64.62	9	61.54	9	64.62	12	69.23
Terminal sacrifice	22	33.85	23	35.38	25	38.46	23	35.38	20	30.77
Total	65		65		65		65		65	

Test	All Dose Groups	Vehicle control vs. Low	Vehicle control vs. Mid	Vehicle control vs. High	Vehicle control vs. Saline
Dose-Response (Likelihood Ratio)	0.6155	0.9342	0.4287	0.7494	0.9325
Homogeneity (Log-Rank)	0.8017	0.9336	0.4234	0.7469	0.9317

#All Cum. % Cumulative Percentage except for Terminal sacrifice;
* = Significant at 5% level; ** = Significant at 1% level.

Table 1B: Intercurrent Mortality Rate in Female Rats

	Vehicle control		Low		Mid		High		Saline Control	
Week / Type of Death	No. of Death	Cum %	No. of Death	Cum %	No. of Death	Cum %	No. of Death	Cum %	No. of Death	Cum %
0 - 52	2	3.08	2	3.08	5	7.69	3	4.62	1	1.54
53 - 78	20	33.85	13	23.08	16	32.31	18	32.31	17	27.69
79 - 91	11	50.77	20	53.85	15	55.38	16	56.92	16	52.31
92 - 104	11	67.69	10	69.23	13	75.38	7	67.69	10	67.69
Terminal sacrifice	21	32.31	20	30.77	16	24.62	21	32.31	21	32.31
Total	65		65		65		65		65	

Test	All Dose Groups	Vehicle control vs. Low	Vehicle control vs. Mid	Vehicle control vs. High	Vehicle control vs. Saline
Dose-Response (Likelihood Ratio)	0.6705	0.9920	0.3510	0.7747	0.9684
Homogeneity (Log-Rank)	0.6966	0.9919	0.3454	0.7720	0.9681

#All Cum. % Cumulative Percentage except for Terminal sacrifice;
* = Significant at 5% level; ** = Significant at 1% level.

Table 2A: Tumor Rates and P-Values for Trend and Pairwise Comparisons in Male Rats

Organ name	Tumor name	Vehicle (VC) 0 mg P - Trend	Low (L) 0.9 mg P - VC vs. L	Mid (M) 3 mg P - VC vs. M	High (H) 9 mg P - VC vs. H	Saline (SC) 0 mg P - VC vs. SC
Body Cavity, Nasal	Adenocarcinoma	0/59 (36)	0/61 (38)	0/62 (41)	1/62 (40)	0/61 (38)
		0.2581	NC	NC	0.5263	NC
	Schwannoma, Malignant	1/59 (37)	0/61 (38)	0/62 (41)	0/62 (39)	0/61 (38)
		1.0000	1.0000	1.0000	1.0000	1.0000
Brain	Astrocytoma, Malignant	0/65 (40)	1/65 (40)	0/65 (43)	2/65 (41)	2/65 (40)
		0.1091	0.5000	NC	0.2531	0.2468
	Oligodendroglioma, Malignant	1/65 (40)	1/65 (40)	0/65 (43)	0/65 (41)	0/65 (40)
		0.9416	NC	1.0000	1.0000	1.0000
Gland, Adrenal	Cortical Adenoma	0/60 (36)	1/61 (38)	1/63 (42)	2/62 (40)	1/63 (40)
		0.1329	0.5135	0.5385	0.2737	0.5263
	Pheochromocytoma, Benign	5/60 (37)	8/61 (40)	7/63 (42)	2/62 (39)	8/63 (41)
		0.9587	0.3261	0.4720	0.9538	0.3443
	Pheochromocytoma, Complex, Benign	0/60 (36)	0/61 (38)	0/63 (42)	1/62 (40)	0/63 (39)
		0.2564	NC	NC	0.5263	NC
	Pheochromocytoma, Benign/ Pheochromocytoma, Complex	5/60 (37)	8/61 (40)	7/63 (42)	3/62 (40)	8/63 (41)
		0.9044	0.3261	0.4720	0.6877	0.6557
	Pheochromocytoma, Malignant	0/60 (36)	1/61 (38)	0/63 (42)	0/62 (39)	0/63 (39)
		0.7677	0.5135	NC	NC	NC
	Pheochromocytoma, Benign/ Pheochromocytoma, Malignant	5/32 (60)	9/31 (61)	7/35 (63)	3/37 (62)	8/33 (63)
		0.9265	0.2349	0.4720	0.6877	0.6557

& X/ZZ (YY): X=number of tumor bearing animals; YY=mortality weighted total number of animals; ZZ=unweighted total number of animals observed;
NC = Not calculable.

Table 2A: Tumor Rates and P-Values for Trend and Pairwise Comparisons in Male Rats (Continued)

Organ name	Tumor name	Vehicle (VC) 0 mg P - Trend	Low (L) 0.9 mg P - VC vs. L	Mid (M) 3 mg P - VC vs. M	High (H) 9 mg P - VC vs. H	Saline (SC) 0 mg P - VC vs. SC
Gland, Mammary	Adenocarcinoma	0/58 (34)	0/60 (37)	0/58 (39)	1/61 (39)	0/59 (36)
		0.2617	NC	NC	0.5342	NC
	Adenoma	0/58 (34)	0/60 (37)	0/58 (39)	1/61 (39)	0/59 (36)
		0.2617	NC	NC	0.5342	NC
	Adenocarcinoma/Adenoma	0/58 (34)	0/60 (37)	0/58 (39)	2/61 (39)	0/59 (36)
		0.0672	NC	NC	0.2820	NC
	Fibroadenoma	2/58 (35)	2/60 (37)	1/58 (39)	1/61 (39)	2/59 (37)
		0.7791	0.7137	0.8990	0.8990	0.7137
	Fibroma	1/58 (34)	0/60 (37)	0/58 (39)	0/61 (39)	0/59 (36)
		1.0000	1.0000	1.0000	1.0000	1.0000
	Fibroadenoma/Fibroma	3/58 (35)	2/60 (37)	1/58 (39)	1/61 (39)	2/59 (37)
		0.8379	0.5268	0.7327	0.7327	0.4732
Gland, Parathyroid	Adenoma	0/55 (32)	0/53 (31)	1/61 (41)	0/54 (34)	1/54 (33)
		0.5435	NC	0.5616	NC	0.5077
Gland, Pituitary	Adenoma	34/64 (53)	34/65 (54)	35/65 (58)	29/61 (51)	40/63 (54)
		0.7930	0.6286	0.7284	0.8317	0.1836
Gland, Prostate	Adenocarcinoma	0/64 (39)	0/65 (40)	0/64 (42)	0/63 (40)	1/63 (40)
		NC	NC	NC	NC	0.5063
	Granular Cell Tumor, Benign	0/64 (39)	0/65 (40)	0/64 (42)	1/63 (41)	0/63 (39)
		0.2531	NC	NC	0.5125	NC
Gland, Thyroid	C-Cell Adenoma	2/55 (35)	4/53 (34)	6/56 (38)	3/57 (38)	3/58 (36)
		0.5535	0.3226	0.1587	0.5396	0.5136
	C-Cell Carcinoma	0/55 (34)	0/53 (34)	0/56 (37)	0/57 (36)	1/58 (36)
		NC	NC	NC	NC	0.5143
	C-Cell Adenoma/ C-Cell Carcinoma	2/55 (35)	4/53 (34)	6/56 (38)	3/57 (38)	4/58 (36)
		0.5304	0.3226	0.1587	0.5396	0.6496
	Follicular Cell Adenoma	1/55 (35)	1/53 (35)	0/56 (37)	3/57 (38)	2/58 (36)
		0.1028	NC	1.0000	0.3391	0.5107

& X/ZZ (YY): X=number of tumor bearing animals; YY=mortality weighted total number of animals; ZZ=unweighted total number of animals observed;
NC = Not calculable.

Table 2A: Tumor Rates and P-Values for Trend and Pairwise Comparisons in Male Rats (Continued)

Organ name	Tumor name	Vehicle (VC) 0 mg P - Trend	Low (L) 0.9 mg P - VC vs. L	Mid (M) 3 mg P - VC vs. M	High (H) 9 mg P - VC vs. H	Saline (SC) 0 mg P - VC vs. SC
Heart	Mesothelioma, Malignant	0/65 (40)	0/65 (40)	0/65 (43)	1/65 (42)	0/65 (40)
		0.2545	NC	NC	0.5122	NC
	Schwannoma, Benign	0/65 (40)	1/65 (40)	0/65 (43)	0/65 (41)	0/65 (40)
		0.7561	0.5000	NC	NC	NC
Hemolymphoreticular Tissue	Histiocytic Sarcoma	0/65 (40)	0/65 (40)	1/65 (44)	2/65 (42)	0/65 (40)
		0.0659	NC	0.5238	0.2593	NC
	Leukemia, Large Granular Lymphocytic	2/65 (40)	0/65 (40)	2/65 (43)	0/65 (41)	1/65 (40)
		0.8239	1.0000	0.7191	1.0000	0.8797
	Lymphoma, Malignant	1/65 (41)	0/65 (40)	2/65 (44)	1/65 (42)	1/65 (41)
		0.3837	1.0000	0.5268	0.7590	NC
Kidney	Lipoma	0/60 (37)	2/60 (37)	0/63 (41)	0/65 (41)	0/64 (40)
		0.8317	0.2466	NC	NC	NC
Large Intestine, Cecum	Fibroma	0/53 (33)	0/50 (32)	1/54 (37)	0/56 (36)	0/54 (35)
		0.5290	NC	0.5286	NC	NC
Lung	Bronchioloalveolar Adenoma	0/59 (37)	0/56 (35)	1/61 (41)	1/62 (39)	0/62 (39)
		0.2039	NC	0.5256	0.5132	NC
Lymph Node, Mesenteric	Hemangioma	0/59 (36)	2/60 (39)	1/62 (41)	1/63 (40)	2/62 (38)
		0.4441	0.2670	0.5325	0.5263	0.2603
Pancreas	Acinar-Islet Cell Adenoma	0/62 (38)	0/62 (39)	1/64 (43)	1/64 (40)	0/63 (39)
		0.1965	NC	0.5309	0.5128	NC
	Adenoma	1/62 (38)	1/62 (39)	0/64 (42)	1/64 (40)	0/63 (39)
		0.5407	0.7597	1.0000	0.7659	1.0000
	Islet Cell Adenoma	0/62 (38)	5/62 (40)	5/64 (43)	2/64 (40)	1/63 (39)
		0.5520	0.0312 $	0.0376 $	0.2597	0.5065
	Islet Cell Carcinoma	1/62 (38)	0/62 (39)	0/64 (42)	1/64 (41)	0/63 (39)
		0.4480	1.0000	1.0000	0.7718	1.0000
	Islet Cell Adenoma/ Islet Cell Carcinoma	1/62 (38)	5/62 (40)	5/64 (43)	3/64 (41)	1/63 (39)
		0.4874	0.1123	0.1315	0.3370	0.7468

& X/ZZ (YY): X=number of tumor bearing animals; YY=mortality weighted total number of animals; ZZ=unweighted total number of animals observed;
NC = Not calculable.

Table 2A: Tumor Rates and P-Values for Trend and Pairwise Comparisons in Male Rats (Continued)

Organ name	Tumor name	Vehicle (VC) 0 mg P - Trend	Low (L) 0.9 mg P - VC vs. L	Mid (M) 3 mg P - VC vs. M	High (H) 9 mg P - VC vs. H	Saline (SC) 0 mg P - VC vs. SC
Skin	Basal Cell Tumor, Benign	1/65 (40)	0/65 (40)	0/62 (41)	0/65 (41)	0/65 (40)
		1.0000	1.0000	1.0000	1.0000	1.0000
	Fibroma	5/65 (40)	2/65 (40)	6/62 (43)	5/65 (42)	6/65 (42)
		0.3410	0.9458	0.5521	0.6624	0.5352
	Fibrosarcoma	1/65 (40)	0/65 (40)	1/62 (42)	3/65 (42)	0/65 (40)
		0.0677	1.0000	0.7651	0.3265	1.0000
	Fibroma/Fibrosarcoma	6/65 (41)	2/65 (40)	7/62 (43)	8/65 (43)	6/65 (42)
		0.1191	0.8600	0.5377	0.4236	0.6044
	Fibrous Histiocytoma, Malignant	0/65 (40)	2/65 (40)	0/62 (41)	0/65 (41)	0/65 (40)
		0.8175	0.2468	NC	NC	NC
	Keratoacanthoma	1/65 (40)	3/65 (40)	0/62 (41)	1/65 (41)	3/65 (41)
		0.7312	0.3077	1.0000	0.7593	0.3172
	Squamous Cell Carcinoma	0/65 (40)	0/65 (40)	0/62 (41)	0/65 (41)	1/65 (40)
		NC	NC	NC	NC	0.5000
	Keratoacanthoma/ Squamous Cell Carcinoma	1/65 (40)	3/65 (40)	0/62 (41)	1/65 (41)	4/65 (41)
		0.7127	0.3077	0.5062	0.2531	0.8127
	Lipoma	1/65 (40)	5/65 (41)	5/62 (42)	0/65 (41)	3/65 (40)
		0.9349	0.1062	0.1122	1.0000	0.3077
	Papilloma	0/65 (40)	0/65 (40)	0/62 (41)	1/65 (41)	0/65 (40)
		0.2531	NC	NC	0.5062	NC
	Sarcoma	1/65 (40)	0/65 (40)	1/62 (42)	1/65 (42)	1/65 (40)
		0.4071	1.0000	0.7651	0.7651	NC
	Schwannoma, Benign	1/65 (40)	0/65 (40)	0/62 (41)	0/65 (41)	0/65 (40)
		1.0000	1.0000	1.0000	1.0000	1.0000
Small Intestine, Duodenum	Adenocarcinoma	0/53 (33)	0/54 (34)	0/58 (38)	0/57 (37)	1/56 (36)
		NC	NC	NC	NC	0.5217
Spinal Cord	Astrocytoma, Malignant	0/65 (40)	0/65 (40)	0/65 (43)	0/65 (41)	1/65 (40)
		NC	NC	NC	NC	0.5000

& X/ZZ (YY): X=number of tumor bearing animals; YY=mortality weighted total number of animals; ZZ=unweighted total number of animals observed;
NC = Not calculable.

Table 2A: Tumor Rates and P-Values for Trend and Pairwise Comparisons in Male Rats (Continued)

Organ name	Tumor name	Vehicle (VC) 0 mg P - Trend	Low (L) 0.9 mg P - VC vs. L	Mid (M) 3 mg P - VC vs. M	High (H) 9 mg P - VC vs. H	Saline (SC) 0 mg P - VC vs. SC
Stomach	Leiomyosarcoma	0/58 (36)	0/56 (35)	1/60 (40)	0/62 (40)	0/59 (37)
		0.5298	NC	0.5263	NC	NC
	Squamous Cell Carcinoma	1/58 (36)	0/56 (35)	0/60 (40)	0/62 (40)	0/59 (37)
		1.0000	1.0000	1.0000	1.0000	1.0000
Testis	Interstitial (Leydig) Cell Adenoma	0/65 (40)	1/65 (40)	1/65 (43)	3/65 (43)	1/65 (40)
		0.0491	0.5000	0.5181	0.1343	0.5000
Thymus	Thymoma, Benign	1/39 (24)	0/46 (30)	0/51 (34)	0/45 (28)	0/52 (31)
		1.0000	1.0000	1.0000	1.0000	1.0000
Urinary Bladder	Papilloma	0/63 (39)	0/62 (38)	1/63 (41)	0/63 (40)	0/64 (40)
		0.5127	NC	0.5125	NC	NC

& X/ZZ (YY): X=number of tumor bearing animals; YY=mortality weighted total number of animals; ZZ=unweighted total number of animals observed;
NC = Not calculable.

Table 2B: Tumor Rates and P-Values for Trend and Pairwise Comparisons in Female Rats

Organ name	Tumor name	Vehicle (VC) 0 mg P - Trend	Low (L) 0.9 mg P - VC vs. L	Mid (M) 3 mg P - VC vs. M	High (H) 9 mg P - VC vs. H	Saline (SC) 0 mg P - VC vs. SC
Body Cavity, Nasal	Fibrous Histiocytoma, Malignant	0/65 (43)	0/65 (43)	0/64 (39)	0/65 (41)	1/64 (44)
		NC	NC	NC	NC	0.5057
Brain	Astrocytoma, Malignant	0/65 (43)	1/65 (44)	0/65 (40)	0/62 (40)	0/65 (43)
		0.7425	0.5057	NC	NC	NC
	Granular Cell Tumor, Benign	1/65 (43)	0/65 (43)	0/65 (40)	0/62 (40)	0/65 (43)
		1.0000	1.0000	1.0000	1.0000	1.0000
Cervix	Squamous Cell Carcinoma	0/65 (43)	0/65 (43)	0/64 (39)	0/63 (40)	1/65 (43)
		NC	NC	NC	NC	0.5000
Gland, Adrenal	Cortical Adenoma	1/65 (43)	3/63 (43)	3/64 (39)	1/65 (41)	2/65 (43)
		0.6775	0.3080	0.2717	0.7410	0.5000
	Cortical Carcinoma	1/65 (43)	0/63 (42)	0/64 (39)	0/65 (41)	0/65 (43)
		1.0000	1.0000	1.0000	1.0000	1.0000
	Pheochromocytoma, Benign	3/65 (43)	2/63 (43)	0/64 (39)	2/65 (41)	2/65 (43)
		0.5983	0.8200	1.0000	0.8049	0.8200
Gland, Mammary	Adenocarcinoma	12/65 (47)	9/64 (47)	13/64 (43)	13/65 (45)	8/65 (46)
		0.2316	0.8390	0.3964	0.4491	0.8867
	Adenoma	5/65 (44)	10/64 (45)	5/64 (41)	6/65 (43)	5/65 (44)
		0.6084	0.1389	0.5845	0.4835	NC
	Adenocarcinoma/Adenoma	16/65 (47)	16/64 (48)	16/64 (45)	17/65 (47)	13/65 (48)
		0.3915	0.4429	0.5263	0.5000	0.3039
	Adenosquamous Carcinoma	0/65 (43)	0/64 (43)	1/64 (40)	0/65 (41)	0/65 (43)
		0.4850	NC	0.4819	NC	NC
	Carcinosarcoma	0/65 (43)	0/64 (43)	0/64 (39)	1/65 (41)	0/65 (43)
		0.2470	NC	NC	0.4881	NC
	Fibroadenoma	37/65 (54)	32/64 (53)	35/64 (50)	33/65 (52)	36/65 (54)
		0.5944	0.8604	0.5200	0.7747	0.6594
	Fibroma	1/65 (43)	1/64 (43)	0/64 (39)	1/65 (41)	1/65 (43)
		0.5261	NC	1.0000	0.7410	NC
	Fibrosarcoma	0/65 (43)	0/64 (43)	1/64 (39)	0/65 (41)	1/65 (44)
		0.4819	NC	0.4756	NC	0.5057
	Fibroma/Fibrosarcoma	1/65 (43)	1/64 (43)	1/64 (39)	1/65 (41)	2/65 (44)
		0.4975	NC	0.7281	0.7410	0.4913
	Lipoma	0/65 (43)	0/64 (43)	0/64 (39)	1/65 (41)	0/65 (43)
		0.2470	NC	NC	0.4881	NC

& X/ZZ (YY): X=number of tumor bearing animals; YY=mortality weighted total number of animals; ZZ=unweighted total number of animals observed;
NC = Not calculable.

Table 2B: Tumor Rates and P-Values for Trend and Pairwise Comparisons in Female Rats (Continued)

Organ name	Tumor name	Vehicle (VC) 0 mg P - Trend	Low (L) 0.9 mg P - VC vs. L	Mid (M) 3 mg P - VC vs. M	High (H) 9 mg P - VC vs. H	Saline (SC) 0 mg P - VC vs. SC
Gland, Pituitary	Adenoma	43/65 (59)	50/64 (62)	51/65 (59)	41/63 (56)	49/64 (59)
		0.6926	0.2129	0.0541	0.5677	0.1333
	Carcinoma	0/65 (43)	0/64 (43)	1/65 (40)	0/63 (40)	1/64 (43)
		0.4819	NC	0.4819	NC	0.5000
	Adenoma/Carcinoma	43/65 (59)	50/64 (62)	52/65 (60)	41/63 (56)	50/64 (59)
		0.6861	0.2129	0.0494	0.5677	0.9121
Gland, Thyroid	C-Cell Adenoma	1/65 (43)	3/63 (43)	7/63 (40)	2/65 (41)	3/64 (44)
		0.4777	0.3080	0.0223	0.4819	0.3168
	C-Cell Carcinoma	0/65 (43)	0/63 (43)	0/63 (39)	1/65 (41)	0/64 (43)
		0.2470	NC	NC	0.4881	NC
	C-Cell Adenoma/ C-Cell Carcinoma	1/65 (43)	3/63 (43)	7/63 (40)	3/65 (41)	3/64 (44)
		0.3075	0.3080	0.0223	0.2900	0.6832
	Follicular Cell Adenoma	1/65 (44)	0/63 (43)	1/63 (39)	2/65 (41)	0/64 (43)
		0.1547	1.0000	0.7220	0.4732	1.0000
Heart	Schwannoma, Benign	0/65 (43)	0/65 (43)	1/65 (40)	0/63 (40)	0/65 (43)
		0.4819	NC	0.4819	NC	NC
Hemolymphoretic ular Tissue	Histiocytic Sarcoma	0/65 (43)	1/65 (43)	0/65 (40)	0/65 (41)	0/65 (43)
		0.7425	0.5000	NC	NC	NC
	Leukemia, Large Granular Lymphocytic	2/65 (44)	0/65 (43)	0/65 (40)	1/65 (41)	0/65 (43)
		0.5832	1.0000	1.0000	0.8659	1.0000
	Lymphoma, Malignant	2/65 (44)	0/65 (43)	1/65 (40)	0/65 (41)	0/65 (43)
		0.8634	1.0000	0.8610	1.0000	1.0000
Kidney	Adenoma	1/64 (43)	0/64 (43)	0/63 (39)	0/65 (41)	0/65 (43)
		1.0000	1.0000	1.0000	1.0000	1.0000
	Lipoma	1/64 (43)	1/64 (43)	0/63 (39)	0/65 (41)	0/65 (43)
		0.9341	NC	1.0000	1.0000	1.0000
	Mesenchymal Tumor, Malignant	0/64 (43)	0/64 (43)	0/63 (39)	0/65 (41)	1/65 (44)
		NC	NC	NC	NC	0.5057

& X/ZZ (YY): X=number of tumor bearing animals; YY=mortality weighted total number of animals; ZZ=unweighted total number of animals observed;
NC = Not calculable.

Table 2B: Tumor Rates and P-Values for Trend and Pairwise Comparisons in Female Rats (Continued)

Organ name	Tumor name	Vehicle (VC) 0 mg P - Trend	Low (L) 0.9 mg P - VC vs. L	Mid (M) 3 mg P - VC vs. M	High (H) 9 mg P - VC vs. H	Saline (SC) 0 mg P - VC vs. SC
Liver	Hemangiosarcoma	0/65 (43)	0/65 (43)	0/65 (40)	0/64 (40)	1/65 (44)
		NC	NC	NC	NC	0.5057
	Hepatocellular Adenoma	0/65 (43)	0/65 (43)	1/65 (40)	0/64 (40)	1/65 (44)
		0.4819	NC	0.4819	NC	0.5057
	Hepatocellular Carcinoma	0/65 (43)	0/65 (43)	0/65 (40)	1/64 (40)	0/65 (43)
		0.2410	NC	NC	0.4819	NC
	Hepatocellular Adenoma/ Hepatocellular Carcinoma	0/65 (43)	0/65 (43)	1/65 (40)	1/64 (40)	1/65 (44)
		0.1738	NC	0.4819	0.4819	0.4943
Lung	Bronchioloalveolar Adenoma	0/65 (43)	1/64 (43)	0/63 (39)	0/65 (41)	0/65 (43)
		0.7410	0.5000	NC	NC	NC
Muscle, Skeletal	Hemangioma	0/65 (43)	1/65 (43)	0/65 (40)	0/63 (40)	0/65 (43)
		0.7410	0.5000	NC	NC	NC
Ovary	Granulosa Cell Tumor, Benign	1/64 (43)	0/63 (43)	0/64 (39)	0/63 (40)	0/64 (43)
		1.0000	1.0000	1.0000	1.0000	1.0000
	Luteoma	0/64 (43)	0/63 (43)	0/64 (39)	1/63 (40)	0/64 (43)
		0.2424	NC	NC	0.4819	NC
	Sertoli Cell Tumor, Benign	1/64 (43)	0/63 (43)	0/64 (39)	0/63 (40)	0/64 (43)
		1.0000	1.0000	1.0000	1.0000	1.0000
Pancreas	Islet Cell Adenoma	3/64 (43)	1/65 (44)	0/65 (40)	2/64 (41)	2/65 (44)
		0.5026	0.9446	1.0000	0.8049	0.8270

& X/ZZ (YY): X=number of tumor bearing animals; YY=mortality weighted total number of animals; ZZ=unweighted total number of animals observed;
NC = Not calculable.

Table 2B: Tumor Rates and P-Values for Trend and Pairwise Comparisons in Female Rats (Continued)

Organ name	Tumor name	Vehicle (VC) 0 mg P - Trend	Low (L) 0.9 mg P - VC vs. L	Mid (M) 3 mg P - VC vs. M	High (H) 9 mg P - VC vs. H	Saline (SC) 0 mg P - VC vs. SC
Skin	Fibroma	1/65 (43)	1/65 (44)	2/65 (41)	2/65 (41)	2/65 (43)
		0.2473	0.7586	0.4819	0.4819	0.5000
	Fibrosarcoma	0/65 (43)	1/65 (44)	0/65 (40)	0/65 (41)	0/65 (43)
		0.7440	0.5057	NC	NC	NC
	Fibroma/Fibrosarcoma	1/65 (43)	2/65 (44)	2/65 (41)	2/65 (41)	2/65 (43)
		0.3368	0.5087	0.4819	0.4819	0.5000
	Fibrous Histiocytoma, Benign	0/65 (43)	1/65 (44)	0/65 (40)	0/65 (41)	0/65 (43)
		0.7440	0.5057	NC	NC	NC
	Keratoacanthoma	0/65 (43)	0/65 (43)	2/65 (40)	0/65 (41)	0/65 (43)
		0.4881	NC	0.2292	NC	NC
	Squamous Cell Carcinoma	0/65 (43)	0/65 (43)	0/65 (40)	0/65 (41)	1/65 (44)
		NC	NC	NC	NC	0.5057
	Keratoacanthoma/ Squamous Cell Carcinoma	0/65 (43)	0/65 (43)	2/65 (40)	0/65 (41)	1/65 (44)
		0.4319	NC	0.2292	NC	0.4943
	Lipoma	1/65 (43)	1/65 (44)	0/65 (40)	0/65 (41)	0/65 (43)
		0.9356	0.7586	1.0000	1.0000	1.0000
	Papilloma	0/65 (43)	0/65 (43)	0/65 (40)	0/65 (41)	1/65 (44)
		NC	NC	NC	NC	0.5057
Small Intestine, Duodenum	Leiomyosarcoma	0/63 (42)	1/63 (43)	0/62 (39)	0/64 (40)	0/64 (43)
		0.7439	0.5059	NC	NC	NC
Thymus	Thymoma, Benign	1/59 (39)	1/56 (38)	1/55 (34)	1/54 (35)	0/62 (41)
		0.4808	0.7468	0.7180	0.7257	1.0000
Uterus	Adenocarcinoma	0/65 (43)	1/64 (43)	0/64 (39)	0/63 (40)	0/65 (43)
		0.7394	0.5000	NC	NC	NC
	Endometrial Stromal Polyp	2/65 (44)	0/64 (43)	4/64 (39)	6/63 (40)	4/65 (44)
		0.0103	1.0000	0.2819	0.1039	0.3382

& X/ZZ (YY): X=number of tumor bearing animals; YY=mortality weighted total number of animals; ZZ=unweighted total number of animals observed;
NC = Not calculable.

Table 2B: Tumor Rates and P-Values for Trend and Pairwise Comparisons in Female Rats (Continued)

Organ name	Tumor name	Vehicle (VC) 0 mg P - Trend	Low (L) 0.9 mg P - VC vs. L	Mid (M) 3 mg P - VC vs. M	High (H) 9 mg P - VC vs. H	Saline (SC) 0 mg P - VC vs. SC
Vagina	Granular Cell Tumor, Benign	3/65 (44)	4/65 (45)	2/65 (40)	6/63 (41)	2/64 (42)
		0.1062	0.5127	0.7889	0.2073	0.8051
	Leiomyosarcoma	1/65 (43)	0/65 (43)	0/65 (40)	0/63 (40)	0/64 (42)
		1.0000	1.0000	1.0000	1.0000	1.0000
	Polyp	0/65 (43)	1/65 (43)	0/65 (40)	0/63 (40)	1/64 (42)
		0.7410	0.5000	NC	NC	0.4941
	Sarcoma	0/65 (43)	0/65 (43)	0/65 (40)	0/63 (40)	1/64 (43)
		NC	NC	NC	NC	0.5000

& X/ZZ (YY): X=number of tumor bearing animals; YY=mortality weighted total number of animals; ZZ=unweighted total number of animals observed;
NC = Not calculable.

Table 3A: Intercurrent Mortality Rate in Male Mice

	Vehicle (C)		Low (L)		Mid (M)		High (H)		Saline (N)		Positive (P)	
Week / Type of Death	No. of Death	Cum %	No. of Death	Cum %	No. of Death	Cum %	No. of Death	Cum %	No. of Death	Cum %	No. of Death	Cum %
0 - 13					1	4.00	4	16.00				
14 - 27					2	12.00	3	28.00				
Planned intermittent sacrifice											10	
Terminal sacrifice	25	100.00	25	100.00	22	88.00	18	72.00	25	100.00		
Total	25		25		25		25		25			

Test	All Dose Groups	Vehicle control vs. Low	Vehicle control vs. Mid	Vehicle control vs. High	Vehicle control vs. Saline
Dose-Response (Likelihood Ratio)	0.0003**	NC	0.0384*	0.0011**	NC
Homogeneity (Log-Rank)	0.0014**	NC	0.0770	0.0047**	NC

#All Cum. % Cumulative Percentage except for Terminal sacrifice;
* = Significant at 5% level; ** = Significant at 1% level.

Table 3B: Intercurrent Mortality Rate in Female Mice

	Vehicle (C)		Low (L)		Mid (M)		High (H)		Saline (N)		Positive (P)	
Week / Type of Death	No. of Death	Cum %	No. of Death	Cum %	No. of Death	Cum %	No. of Death	Cum %	No. of Death	Cum %	No. of Death	Cum %
0 - 13	1	4.00	2	8.00			3	12.00				
14 - 27			1	12.00			1	16.00	1	4.00		
Planned intermittent sacrifice											10	
Terminal sacrifice	24	96.00	22	88.00	25	100.00	21	84.00	24	96.00		
Total	25		25		25		25		25			

Test	All Dose Groups	Vehicle control vs. Low	Vehicle control vs. Mid	Vehicle control vs. High	Vehicle control vs. Saline
Dose-Response (Likelihood Ratio)	0.2047	0.3108	0.2390	0.1541	0.9885
Homogeneity (Log-Rank)	0.1500	0.3195	0.3173	0.1676	0.9885

#All Cum. % Cumulative Percentage except for Terminal sacrifice;
* = Significant at 5% level; ** = Significant at 1% level.

Table 4A: Tumor Rates and P-Values for Trend and Pairwise Comparisons in Male Mice

Organ name	Tumor name	Saline (N) P-C vs N	Vehicle (C) 0 mg P-Trend	Low (L) 10 mg P-C vs L	Mid (M) 25 mg P-C vs M	High (H) 75/40 mg P-C vs H	Positive (P) P-C vs P
Cavity, Nasal	Adenocarcinoma	1/25 (25) 0.5000	0/25 (25) NC	0/25 (25) NC	0/25 (23) NC	0/25 (19) NC	
Harderian Glands	Adenoma	1/25 (25) 0.5000	0/25 (25) NC	0/25 (25) NC	0/25 (23) NC	0/25 (19) NC	
	Carcinoma	1/25 (25) 0.5000	0/25 (25) NC	0/25 (25) NC	0/25 (23) NC	0/25 (19) NC	
	Adenoma/Carcinoma	2/23 (25) 0.2449	0/25 (25) NC	0/25 (25) NC	0/23 (25) NC	0/19 (25) NC	
Kidneys	Hemangiosarcoma	0/25 (25) 1.0000	1/25 (25) 0.5700	0/25 (25) 1.0000	1/25 (23) 0.7340	0/25 (19) 1.0000	
Liver	Hepatocellular Adenoma	0/25 (25) 1.0000	1/25 (25) 0.7790	1/25 (25) NC	0/25 (23) 1.0000	0/25 (19) 1.0000	
Lungs With Bronchi	Alveolar-Bronchiolar Adenoma	4/25 (25) NC	4/25 (25) 0.4066	2/25 (25) 0.9053	1/25 (23) 0.9690	3/25 (19) 0.6626	10/10 (10) 0.0000 $
	Alveolar-Bronchiolar Carcinoma	0/25 (25) 1.0000	1/25 (25) 0.7790	1/25 (25) NC	0/25 (23) 1.0000	0/25 (19) 1.0000	0/10 (1) 1.0000
	Alveolar-Bronchiolar Adenoma/ Carcinoma	4/21 (25) 0.5	5/20 (25) 0.5924	3/22 (25) 0.649	1/22 (25) 0.886	3/16 (25) 0.481	10/0 (10) 0.0000 $
Seminal Vesicles	Hemangiosarcoma	0/25 (25) 1.0000	1/25 (25) 0.7283	0/25 (25) 1.0000	0/25 (23) 1.0000	0/25 (19) 1.0000	
Spinal Cord, Lumbar	Hemangiosarcoma	1/25 (25) 0.5000	0/25 (25) NC	0/25 (25) NC	0/25 (23) NC	0/25 (19) NC	
Spleen	Hemangiosarcoma	1/25 (25) 0.5000	0/25 (25) 0.4326	1/25 (25) 0.5000	1/25 (23) 0.4792	0/25 (19) NC	
Thymus	Lymphoma	0/25 (25) NC	0/25 (25) 0.2151	0/25 (25) NC	0/25 (23) NC	1/25 (20) 0.4444	
Whole Body	Hemangiosarcoma	2/25 (25) NC	2/25 (25) 0.8047	1/25 (25) 0.5	2/25 (23) 0.6631	0/25 (19) 0.6829	

& X/ZZ (YY): X=number of tumor bearing animals; YY=mortality weighted total number of animals; ZZ=unweighted total number of animals observed;
NC = Not calculable.

Table 4B: Tumor Rates and P-Values for Trend and Pairwise Comparisons in Female Mice

Organ name	Tumor name	Saline (N) P-C vs N	Vehicle (C) 0 mg P-Trend	Low (L) 10 mg P-C vs L	Mid (M) 25 mg P-C vs M	High (H) 75/40 mg P-C vs H	Positive (P) P-C vs P
Harderian Glands	Adenoma	0/25 (24) 1.0000	1/25 (24) 0.7419	0/25 (23) 1.0000	0/25 (25) 1.0000	0/25 (21) 1.0000	
	Carcinoma	0/25 (24) 1.0000	1/25 (24) 0.7155	1/25 (23) 0.7447	1/25 (25) 0.7653	0/25 (21) 1.0000	
	Adenoma/Carcinoma	0/24 (25) 0.7553	2/22 (25) 0.9246	1/22 (25) 0.4837	1/24 (25) 0.5156	0/21 (25) 0.7212	
Kidneys	Hemangiosarcoma	0/25 (24) NC	0/25 (24) 0.4946	1/25 (23) 0.4894	0/25 (25) NC	0/25 (21) NC	
Lungs With Bronchi	Alveolar-Bronchiolar Adenoma	5/25 (24) 0.0941	1/25 (24) 0.8152	2/25 (23) 0.4837	1/25 (25) 0.7653	0/25 (21) 1.0000	10/10 (10) 0.0000 $
	Alveolar-Bronchiolar Carcinoma	5/19 (25) 0.0941	1/23 (25) 0.8321	2/21 (25) 0.4837	2/23 (25) 0.5156	0/21 (25) 0.4667	10/0 (10) 0.0000 $
	Alveolar-Bronchiolar Adenoma/ Carcinoma	0/25 (24) NC	0/25 (24) 0.4946	0/25 (23) NC	1/25 (25) 0.5102	0/25 (21) NC	1/10 (2) 0.0769
	Hemangiosarcoma	1/25 (25) 0.5102	0/25 (24) NC	0/25 (23) NC	0/25 (25) NC	0/25 (21) NC	0/10 (1) NC
Mammary Gland	Lymphangioma	0/25 (24) NC	0/25 (24) 0.2258	0/25 (23) NC	0/25 (25) NC	1/25 (21) 0.4667	
Multicentric	Mesothelioma	1/25 (25) 0.5102	0/25 (24) NC	0/25 (23) NC	0/25 (25) NC	0/25 (21) NC	
Skin - Soi (Standard Section)	Squamous Cell Carcinoma	0/25 (24) 1.0000	1/25 (24) 0.7419	0/25 (23) 1.0000	0/25 (25) 1.0000	0/25 (21) 1.0000	
Spleen	Hemangiosarcoma	1/25 (24) 0.5000	0/25 (24) NC	0/25 (23) NC	0/25 (25) NC	0/25 (21) NC	
Stomach	Squamous Cell Carcinoma	0/25 (24) NC	0/25 (24) 0.4946	1/25 (23) 0.4894	0/25 (25) NC	0/25 (21) NC	
Thymus	Thymoma	0/25 (24) NC	0/25 (24) 0.4727	1/25 (23) 0.4894	1/25 (25) 0.5102	0/25 (21) NC	
Uterus	Endometrial Stromal Polyp	0/25 (24) 1.0000	1/25 (24) 0.7419	0/25 (23) 1.0000	0/25 (25) 1.0000	0/25 (21) 1.0000	
Whole Body	Hemangiosarcoma	2/25 (25) 0.2551	0/25 (24) 0.4946	1/25 (23) 0.4894	0/25 (25) NC	0/25 (21) NC	

& X/ZZ (YY): X=number of tumor bearing animals; YY=mortality weighted total number of animals; ZZ=unweighted total number of animals observed;
NC = Not calculable.

Figure 1A: Kaplan-Meier Survival Functions for Male Rats

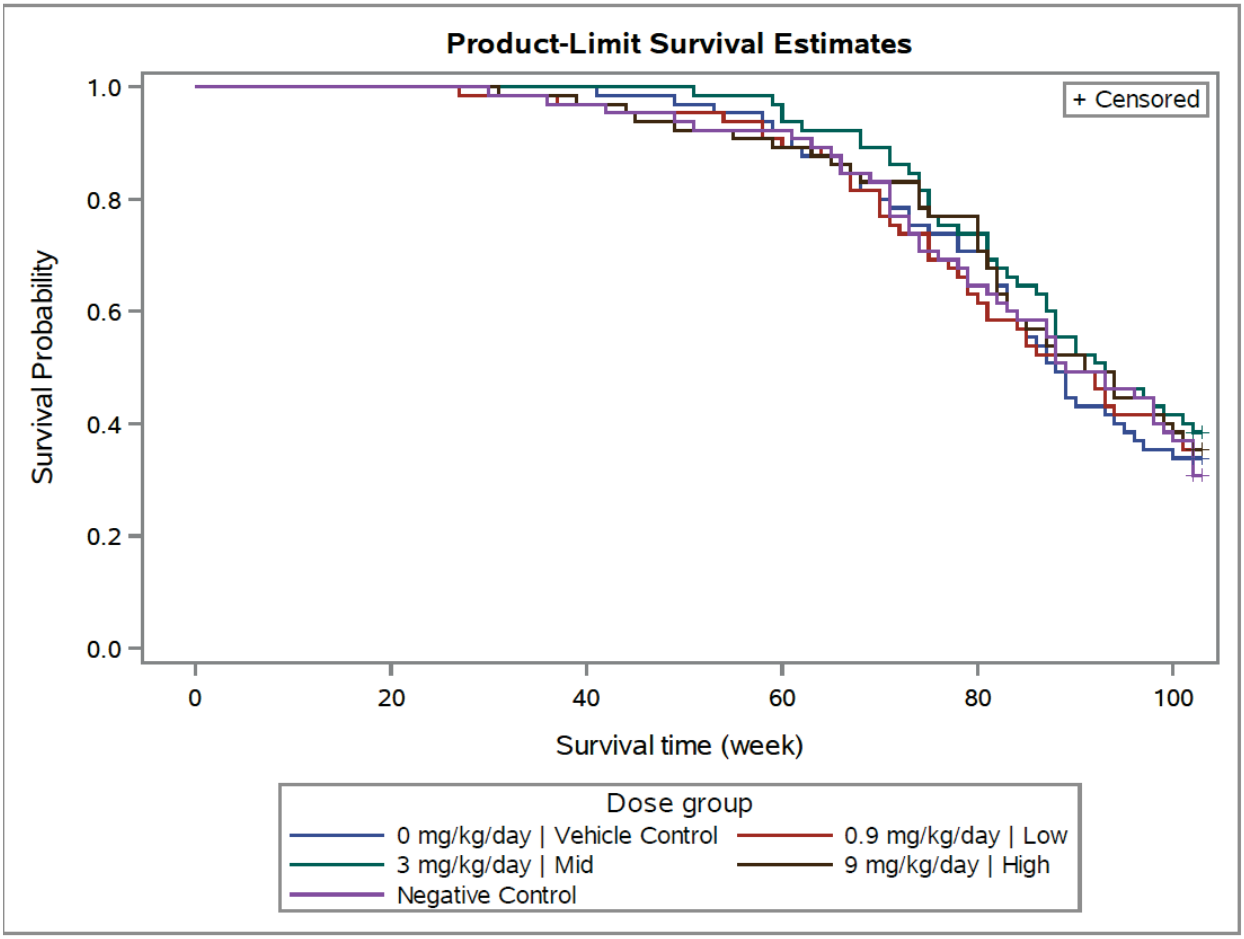


Figure 1B: Kaplan-Meier Survival Functions for Female Rats

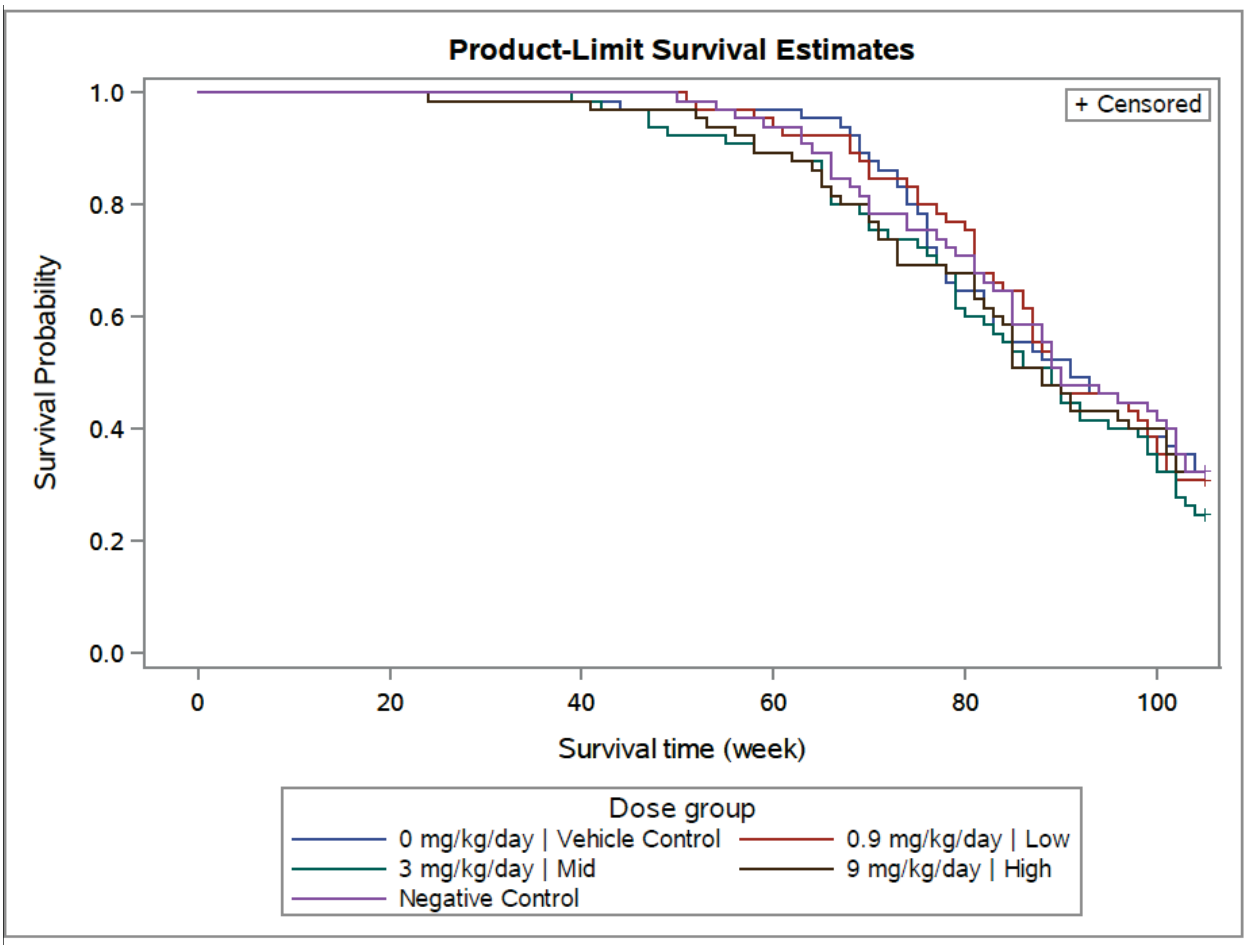


Figure 2A: Kaplan-Meier Survival Functions for Male Mice

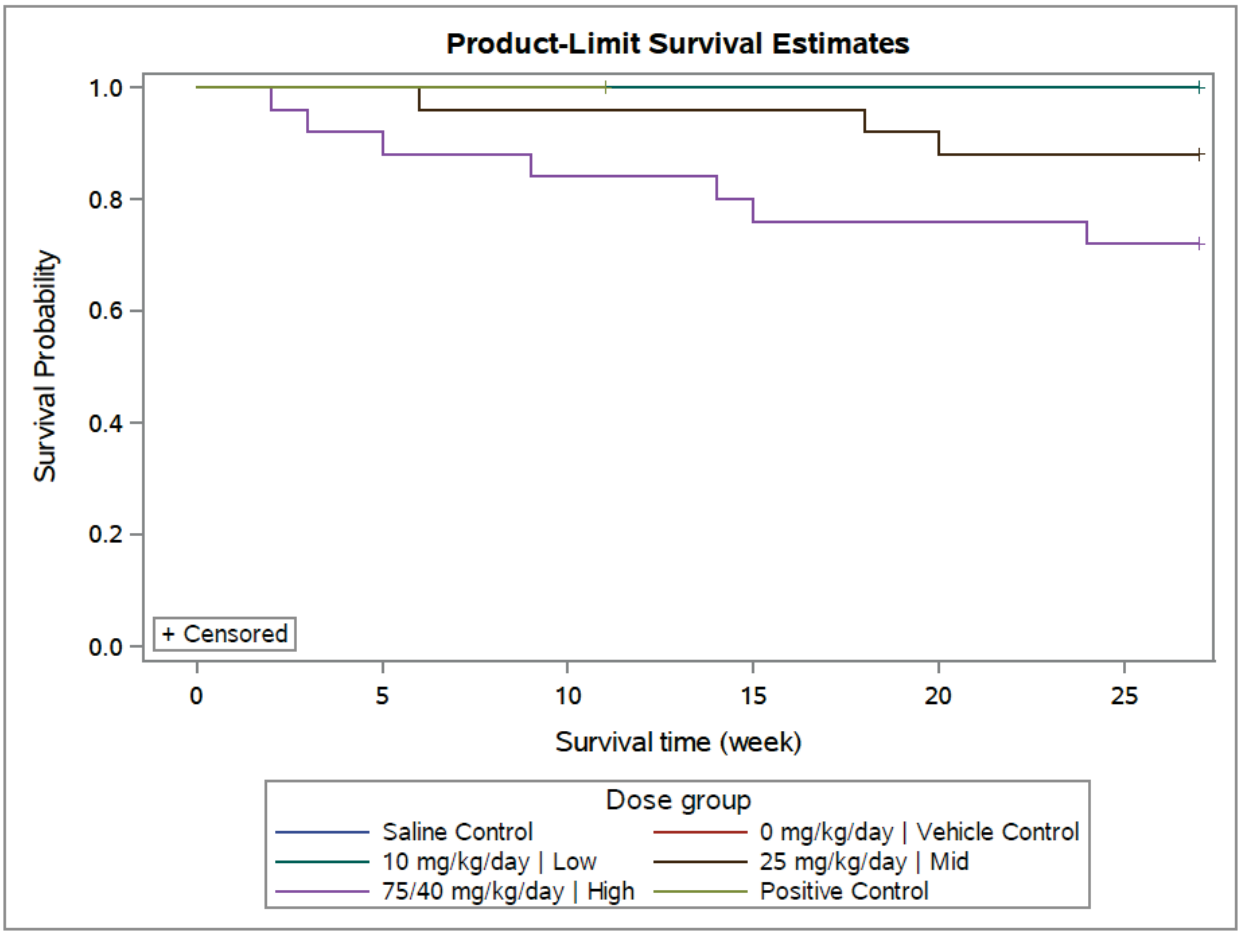
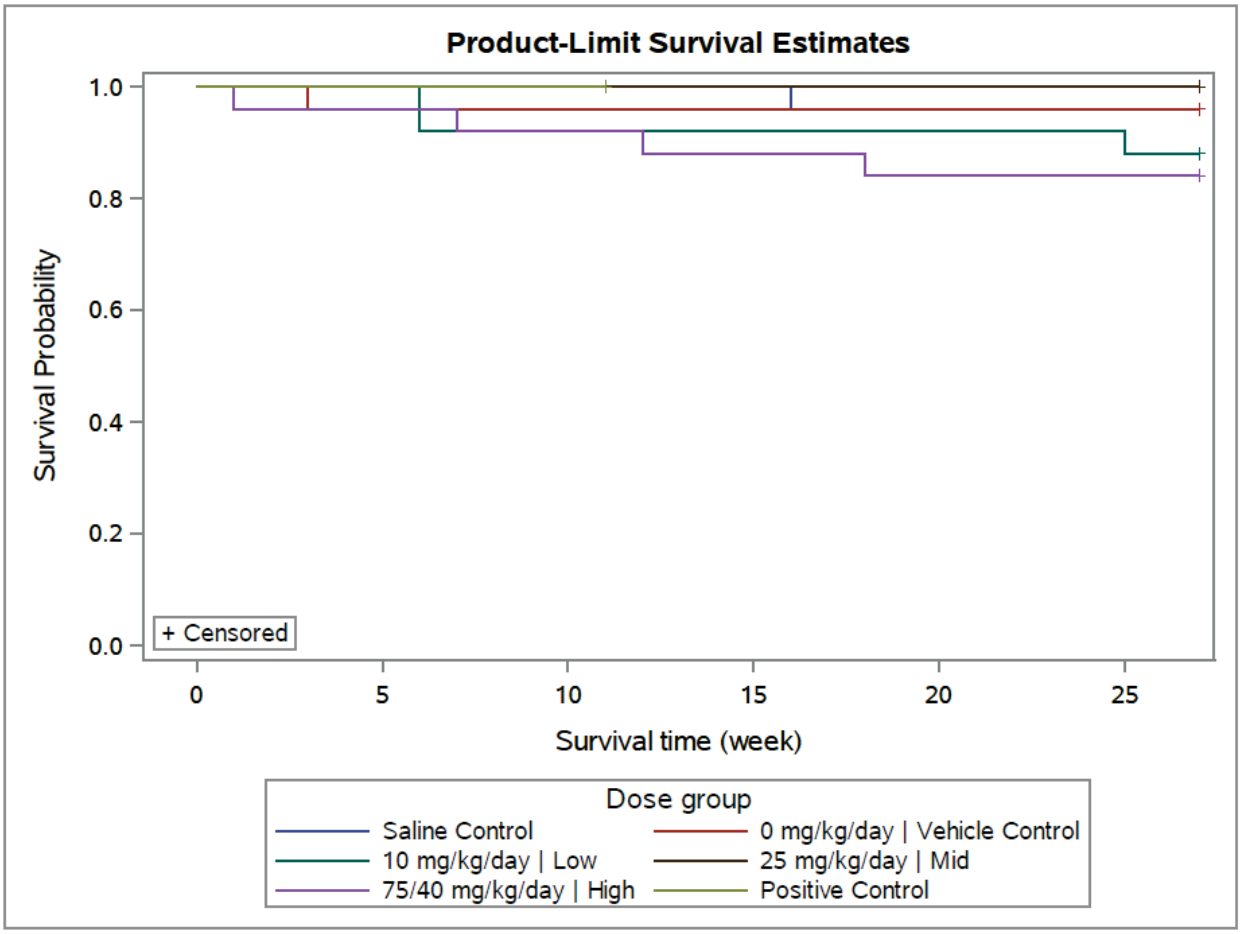


Figure 2B: Kaplan-Meier Survival Functions for Female Mice



6. References

Bailer, A.J, Portier, C.J. (1988). “Effects of treatment-induced mortality and tumor-induced mortality on tests for carcinogenicity in small samples.” *Biometrics*, 44, 417-431.

Bieler, G.S. and Williams, R.L. (1993). “Ratio estimates, the delta method, and quantal response tests for increased carcinogenicity”. *Biometrics* 49, 793-801.

Guidance for Industry. Statistical Aspects of the Design, Analysis, and Interpretation of Chronic Rodent Carcinogenicity Statues of Pharmaceuticals (Draft Guidance). U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER). May 2001.

Mantel N. (1980). Assessing laboratory evidence for neoplastic activity. *Biometrics*. 36: 381-399.

Heyse J.F., Rom D. (1988). Adjusting for Multiplicity of Statistical Tests in the Analysis of Carcinogenicity Studies. *Biometrical Journal*. 30: 883-896.

Harter H.I. (1957). Error Rates and Sample Sizes for Range Tests in Multiple Comparisons. *Biomtrics*. 13: 511-536.

Haseman, J. (1983). “A re-examination of false-positive rates for carcinogenesis studies”, *Fundamental and Applied Toxicology*, 3: 334-339.

Lin KK (2000) Carcinogenicity Studies of Pharmaceuticals. In: *Encyclopedia of Biopharmaceutical Statistics*, ed. Shein-Chung Chow, Marcel Dekker, New York.

Lin K.K. and Rahman A.M. (1998). ” Overall false positive rates in tests for linear trend in tumor incidence in animal carcinogenicity studies of new drugs”, *Journal of Biopharmaceutical Statistics*, 8(1), 1-15.

Peto, R., M.C. Pike, N.E. Day, R.G. Gray, P.N. Lee, S. Parish, J. Peto, Richards, and J.Wahrendorf. (1980) “Guidelines for sample sensitive significance test for carcinogenic effects in long-term animal experiments”, Long term and short term screening assays for carcinogens: A critical appraisal, International agency for research against cancer monographs, *Annex to supplement, World Health Organization, Geneva*, 311-426.

Rahman, A.M., and Lin, K.K. (2008), "A Comparison of False Positive Rates of Peto and Poly-3 Methods for Long-Term Carcinogenicity Data Analysis Using Multiple Comparison Adjustment Method Suggested by Lin and Rahman", *Journal of Biopharmaceutical Statistics*, 18:5, 849-858.

Rahman, A.M., and Lin, K.K. (2009), "Design and Analysis of Chronic Carcinogenicity Studies of Pharmaceuticals in Rodents", in “Design and Analysis of Clinical Trials with Time-to-Event Endpoints", K.E Peace, Editor, Chapman & Hall/CRC, Taylor & Francis Group, LLC, Boca Raton, FL, London, and New York.

This is a representation of an electronic record that was signed electronically. Following this are manifestations of any and all electronic signatures for this electronic record.

/s/

HEPEI CHEN
11/21/2018

KARL K LIN
11/21/2018
concur with review.