The Effectiveness of Esketamine on Depression Alleviation Based on Mixed Model for Repeated Measures

Abstract

Background: Depression represents a significant global health burden, prompting increased research into novel therapeutic interventions. Esketamine has emerged as a promising treatment for patients with treatment-resistant depression.

Objective: This study evaluates the efficacy of Esketamine at two different dosages (56 mg and 84 mg) compared to placebo in patients with major depressive disorder.

Methods: In this randomized controlled trial, patients were allocated in a 1:1:2 ratio to Esketamine 56 mg, Esketamine 84 mg, or placebo groups. Treatment outcomes were assessed using Mixed Models for Repeated Measures (MMRM) with an unstructured covariance matrix to account for longitudinal measurements.

Results: Statistical analysis using the least squares means estimation (from the emmeans package in R) revealed significant differences in depression severity scores between both Esketamine dosage groups and placebo. Both 56 mg and 84 mg Esketamine demonstrated clinically meaningful improvement in depressive symptoms compared to placebo across assessment timepoints.

Conclusion: Esketamine represents an effective pharmacological intervention for depression, with both studied dosages showing superior efficacy compared to placebo. These findings support Esketamine as a valuable addition to the therapeutic arsenal for patients suffering from depression.

Keywords: Depression; Esketamine; MMRM; Emmeans

1 Introduction

Depression has become a significant global health concern, especially in recent years, as highlighted by the World Health Organization (WHO). It is one of the leading causes of disability worldwide, affecting millions of individuals across different age groups, genders, and socioeconomic backgrounds. The increasing prevalence of

depression has drawn urgent attention from researchers, clinicians, and policymakers due to its profound impact on individuals' quality of life, productivity, and overall well-being. As a major psychological health issue, depression poses a serious threat to human health and has far-reaching social and economic consequences. Addressing this challenge requires effective therapeutic interventions and robust clinical evidence to guide treatment strategies.

In response to the growing burden of depression, numerous clinical trials have been conducted to explore innovative treatments. Among these, Esketamine, a derivative of ketamine, has emerged as a promising therapeutic option for treatment-resistant depression. Esketamine has been shown to act rapidly on the glutamatergic system, offering a novel mechanism of action compared to traditional antidepressants that primarily target monoaminergic pathways. Clinical studies have demonstrated its potential to reduce depressive symptoms in patients who have not responded adequately to conventional therapies. However, further research is needed to evaluate its efficacy across different dosages and patient subgroups.

This study utilizes the Mixed Model for Repeated Measures (MMRM) to analyze the effectiveness of Esketamine in alleviating depression, leveraging its advantages over traditional linear regression models. Unlike linear regression, which assumes independence of observations, the MMRM model accounts for the correlation between repeated measurements within the same subject over time, making it particularly suitable for longitudinal clinical trial data with outcomes measured across multiple time points. Using clinical trial data provided by Tigermed, the study evaluates the impact of two dosages of Esketamine (56 mg and 84 mg) on Montgomery–Åsberg Depression Rating Scale (MADRS) scores, a widely recognized measure of depression severity. The randomized, double-blind, placebo-controlled trial assigned participants to one of three groups—Esketamine 56 mg, Esketamine 84 mg, or placebo—in a 1:1:2 ratio. Through the application of the MMRM model and the estimated marginal means (EMMeans) method, the study compares treatment effects between the placebo group and the two Esketamine groups to provide robust insights into its therapeutic efficacy. Furthermore, this study explores the robustness of Esketamine's therapeutic effects through subgroup analyses and sensitivity analyses. Specifically, we examine its efficacy across different age groups (19–60 years and >60 years) and among patients with moderate and severe depression. The results indicate that Esketamine demonstrates consistent effectiveness across these subgroups, suggesting its potential as a versatile treatment option for diverse patient populations.

The remainder of this paper is organized as follows. Section 2 introduces the data source and outlines the criteria for participant inclusion in the study. Section 3 presents the structure of the MMRM model, details of the EMMeans analysis, and the methodological framework employed in this study. Section 3 provides the main results, followed by an extensive discussion in Section 5. Finally, Section 6 concludes the paper with a summary of findings and implications for future research.

Main results are presented in Section 4 followed by an extensive discussion. We conclude this paper in Section 6.

2 Data

This study enrolled 474 participants from the United States who voluntarily participated.

Inclusion criteria.

Adults aged 18 years or older who experienced first onset of depressive symptoms before age 55

Met the DSM-5 diagnostic criteria for either: Single-episode Major Depressive Disorder (MDD) with a minimum episode duration of ≥2 years, or Recurrent MDD without psychotic features

Demonstrated treatment resistance, defined as nonresponse (≤25% improvement) to ≥2 oral antidepressant treatments during the current depressive episode

Scored ≥ 34 on the Inventory of Depressive Symptomatology–Clinician Rated ($IDS - C_{30}$) total score

Were medically stable as determined by physical examination

Exclusion criteria.

Participants were excluded from the study if they met any of the following criteria: prior use of ketamine or Esketamine (lifetime); previous nonresponse to an adequate course of electroconvulsive therapy (ECT) in the current major depressive episode, defined as at least 7 treatments with unilateral or bilateral ECT; receipt of vagal nerve stimulation (VNS) or deep brain stimulation (DBS) in the current depressive episode; current or prior DSM-5 diagnosis of a psychotic disorder or Major Depressive Disorder with psychotic features, bipolar or related disorders (confirmed by the Mini-International Neuropsychiatric Interview [MINI]), current obsessive-compulsive disorder, intellectual disability (DSM-5 diagnostic codes 317, 318.0, 318.1, 318.2, 315.8, and 319), autism spectrum disorder, or personality disorders including borderline, antisocial, histrionic, or narcissistic types.

Data Selection and Variables.

In selecting the clinical data records for participants, we exclusively utilized data from the double-blind phase (primarily covering days 0-28). For the critical Montgomery-Åsberg Depression Rating Scale (MADRS) assessments, we specifically selected the Derived MADRS Total Score(i.e: MADA0212 socre) while for temporal variables, we utilized only the Numeric representation of analysis visit (AVISITN). The key variables in this study were factors for treatment (TRT01A) and specific days within the double-blind phase (AVISITN), as well as their interaction effects. The primary efficacy variable was the MADR0212 score (AVAL), with higher scores indicating greater severity of depressive symptoms. Consequently, a greater reduction in scores represents superior therapeutic efficacy. Additional covariates included age (AGE), study center (SITEID), MADRS baseline score (BASE), and race (RACE).

3 Methodology

3.1 Flow Chart

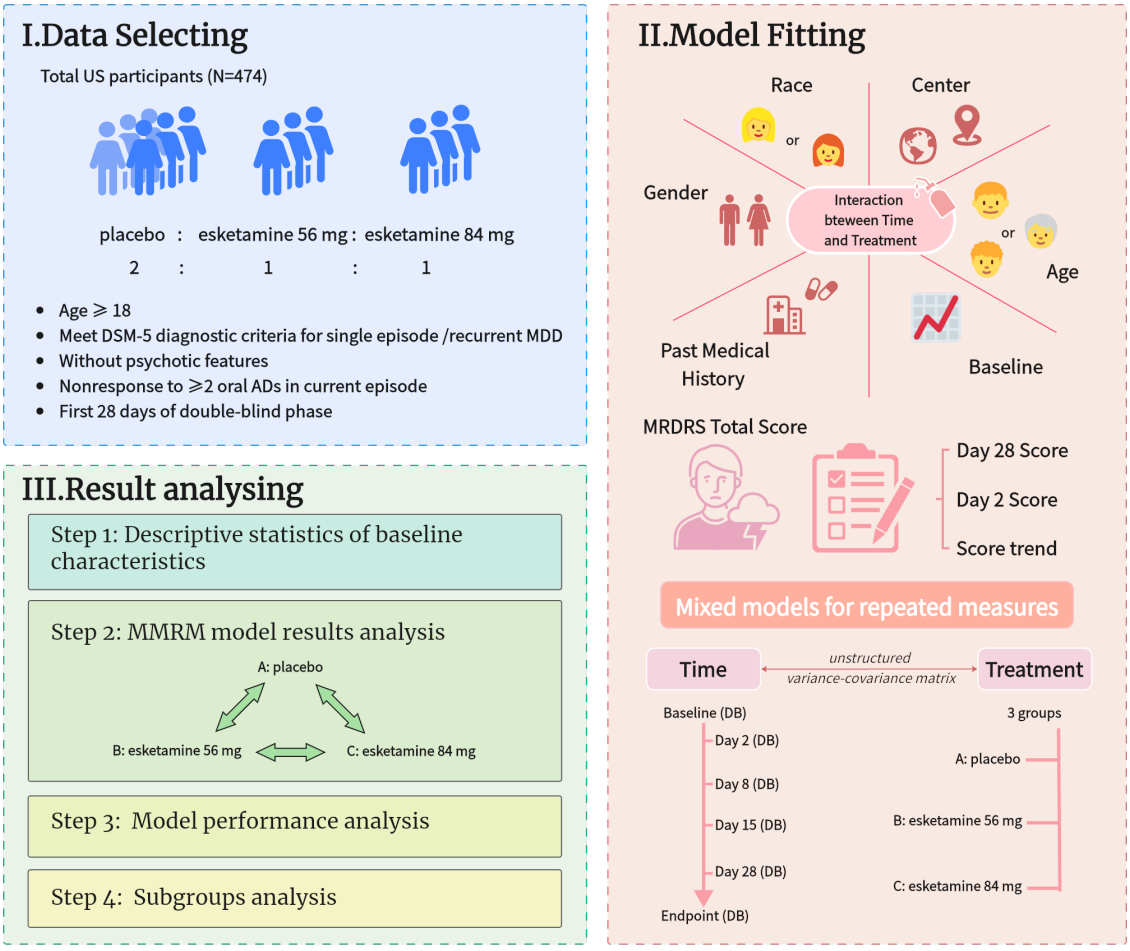

Fig. 1: Work Flow Chart

3.2 Model Structure

The Mixed Model for Repeated Measures (MMRM) proposed by Sabanes (2025) is a statistical model designed to analyze longitudinal sequential results, particularly in randomized clinical trials and other scenarios requiring the processing of repeated measurements. This model extends the basic linear mixed-effects model introduced by Laird and Ware (1982).

3.2.1 The Basic Linear Mixed-Effects Model

Firstly, we discuss the basic linear mixed-effects model:

$$y_i = X_i\beta + Z_i b_i + \epsilon_i, \quad i = 1, \ldots, M \tag{1}$$

where

- y_i is the n_i-dimensional response vector for the ith subject.
- β is the p-dimensional vector of fixed effects, which are the parameters of interest to the researchers.
- b_i is the q-dimensional vector of random patient-specific effects.
- X_i (of size $n_i \times p$) and Z_i (of size $n_i \times q$) are known regressor matrices relating observations to the fixed-effects and random-effects, respectively.
- ϵ_i is the n_i-dimensional within-subject error.

This model assumes that:

- b_i and ϵ_i follow a normal distribution with a mean of 0, and variance-covariance matrices are Ψ and $\sigma^2 I$, respectively.
- $b_i \sim \mathcal{N}(0, \Psi)$ and $\epsilon_i \sim \mathcal{N}(0, \sigma^2 I)$.
- b_i and ϵ_i are independent for different subjects and independent of each other for the same subject.

3.2.2 Mixed Model for Repeated Measures (MMRM)

Now we assume that Ψ and σ^2 are different for different individuals, i.e.,

$$\epsilon_i \sim \mathcal{N}(0, \Lambda_i), \quad i = 1, \ldots, M. \tag{2}$$

where the Λ_i are positive-definite matrices parameterized by a fixed, generally small set of parameters λ. Similarly, b_i and ϵ_i are independent for different subjects and independent of each other for the same subject. The variance-covariance matrix of the response vector y_i,

$$\text{Var}(y_i) = \Sigma_i = \left(Z_i \Psi Z_i^T + \Lambda_i \right) \tag{3}$$

comprises a random-effects component, given by $Z_i \Psi Z_i^T$, and a within-subject component, given by Λ_i. The MMRM is a special case of the previous equation. In a clinical trial setting, one often chooses to directly model the variance-covariance structure of the response, i.e., to account for within-subject dependencies using the within-group component Λ_i, and can omit the random effects component ($Z_i b_i$). Hence, in this case,

$$\text{Var}(y_i) = \Sigma_i = \Lambda_i \tag{4}$$

Now we get the MMRM:

$$y_i = X_i \beta + \epsilon_i, \quad \epsilon_i \sim \mathcal{N}(0, \Sigma_i), \quad i = 1, \ldots, M. \tag{5}$$

The Σ_i matrices are obtained by subsetting the overall variance-covariance matrix $\Sigma \in \mathbb{R}^{m \times m}$, where m is the total number of scheduled visits per subject, appropriately by

$$\Sigma_i = S_i^\top \Sigma S_i \tag{6}$$

where $S_i \in \{0, 1\}^{m \times m_i}$ is the subject-specific "subsetting matrix" that indicates the visits with available observations.

In our application:

- β is a vector of treatments (dummy A, B, C), time (Day 0, 2, 8, 15, 22, 28, Endpoint), the interaction of treatments and time, SITEID, RACE, and AGE.
- X_i is a designed matrix (474*6) related to fixed effect β.
- y_i is AVAL for each subject (USUBJID) at each AVISITN.
- ϵ_i is an unstructured covariance matrix.

3.2.3 Model Advantage

The Mixed Model for Repeated Measures (MMRM) is a powerful statistical tool particularly well-suited for analyzing longitudinal data, which involves multiple measurements taken on the same subjects over time. By accounting for within-subject correlations, MMRM effectively handles repeated measurements and produces more accurate analytical results. It can also make full use of all available data, performing likelihood-based estimation to conduct valid inference even when some observations are missing. By incorporating random effects, MMRM captures the variability among individuals and enhances the flexibility of the model. Moreover, it supports a variety of covariance structures, such as unstructured and autoregressive, allowing for flexible modeling of within-subject correlations. Through proper modeling, MMRM can provide more accurate parameter estimates and statistical inferences. Widely applied in clinical trials and longitudinal studies, MMRM can effectively assess intervention effects and temporal trends. In this study, we utilized the MMRM model to analyze the longitudinal data of AVAL for the same USUBJID, taking into account repeated measurements, individual differences, and flexible covariance structures, thereby providing strong statistical support for our research questions.

3.2.4 Emeans

This study employed estimated marginal means (emmeans) to calculate adjusted means from the Mixed Model Repeated Measures (MMRM) analysis. The core objective of the emmeans algorithm is to provide more accurate between-group comparisons by adjusting for covariates or balancing the effects of experimental design. Marginal means, also known as least-squares means, are group means estimated through statistical modeling that adjust for the influence of other variables. The calculation process involves first constructing a reference grid, where factors and covariates in the model are fixed at specific values (covariates set to sample means, factors at all levels). Predicted means are then calculated based on model coefficients for each factor level under the reference grid conditions. Finally, averaging procedures are applied to eliminate the interference of covariates in between-group comparisons.

The emmeans approach offers advantages in flexibility, interpretability, and extensibility by constructing reference grids and model predictions for comparing group means while adjusting for covariates. Given the multiple covariates involved in this study, the emmeans algorithm was necessary to eliminate covariate interference before performing between-group comparisons. Additionally, since treatment groups and time

points showed interaction effects in this study, emmeans helped estimate means for different groups at various time points, demonstrating the efficacy trends of Esketamine nasal spray.

The study specified the interaction between treatment group (TRT01A) and visit number (AVISITN) as the factors of interest for comparison. Due to the predominance of white participants in the study population and the relatively small numbers of participants from other racial groups (limiting representativeness and research significance), race (RACE) was set as an irrelevant variable. In the actual analysis, the model calculated estimated marginal means using the emmeans function and adjusted for confounding effects of baseline scores, visit time, study center (SITEID), and baseline symptoms (BaseSituation). The p-values for data comparisons underwent Dunnett's test twice to ensure statistical significance and control for Type I error in multiple comparisons.

4 Results

4.1 Descriptive Statistics Analysis

In our analysis, we categorized patients into two groups based on their MADRS scores: those with scores less than 40, indicating middle-level depression, and those with scores of 40 or higher, indicating serious-level depression. The table provides a detailed breakdown of various demographic and clinical characteristics across these groups.

The average age of the total sample was 45.2 years, with those in the middle-level depression group slightly younger at 45.5 years compared to 44.4 years in the serious-level depression group. This difference was statistically significant (p=0.045). The majority of patients in both groups were female, with a higher proportion in the middle-level depression group (58.2%)compared to the serious-level depression group (64.0%). This difference was also statistically significant (p=0.006). The largest racial group in both categories was White, with 86.0% in the middle-level depression group and 85.3% in the serious-level depression group. Other racial categories showed no significant differences between the groups. TRT01A: There was no significant difference in the distribution of treatment types (Dummy A, B, and C) between the two groups (p=0.429). The distribution of baseline severity groups (Group1 to Group5) showed significant differences between the two depression levels. Group1, which represents the least severe depression, was more prevalent in the middle-level depression group (19.2%) compared to the serious-level depression group (22.8%). Conversely, Group5,

representing the most severe depression, was more common in the serious-level depression group (14.9%) compared to the middle-level depression group (21.4%) (p<0.001).

This analysis provides a comprehensive overview of the baseline characteristics of patients with varying levels of depression as indicated by table 1, highlighting significant differences in age, sex, and baseline severity distribution between middle-level and serious-level depression groups. The specific information of 3 treatments situation in different sites are shown in the figures 2 , 3 and 4.

	[ALL] N=3241	**MADRS SCORE < 40** N=2525	**MADRS SCORE ≥ 40** N=716	**p.overall**
AGE	45.2 (14.0)	45.5 (14.5)	44.4 (12.3)	0.045
SEX:				0.006
F	1928 (59.5%)	1470 (58.2%)	458 (64.0%)	
M	1313 (40.5%)	1055 (41.8%)	258 (36.0%)	
RACE:				
WHITE	2788 (86.0%)	2177 (86.2%)	611 (85.3%)	
AMERICAN INDIAN OR ALASKA NATIVE	6 (0.19%)	6 (0.24%)	0 (0.00%)	
ASIAN	97 (2.99%)	83 (3.29%)	14 (1.96%)	
BLACK OR AFRICAN AMERICAN	241 (7.44%)	185 (7.33%)	56 (7.82%)	
MULTIPLE	60 (1.85%)	46 (1.82%)	14 (1.96%)	
NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER	14 (0.43%)	7 (0.28%)	7 (0.98%)	
NOT REPORTED	28 (0.86%)	21 (0.83%)	7 (0.98%)	
UNKNOWN	7 (0.22%)	0 (0.00%)	7 (0.98%)	
TRT01A:				0.429
Dummy A	1610 (49.7%)	1252 (49.6%)	358 (50.0%)	
Dummy B	819 (25.3%)	650 (25.7%)	169 (23.6%)	
Dummy C	812 (25.1%)	623 (24.7%)	189 (26.4%)	
BASE	35.4 (6.36)	33.1 (5.09)	43.6 (2.26)	0.000
AVAL	27.9 (11.1)	26.0 (10.0)	34.5 (12.1)	<0.001
SITEID_group:				<0.001
Group1	649 (20.0%)	486 (19.2%)	163 (22.8%)	
Group2	649 (20.0%)	516 (20.4%)	133 (18.6%)	
Group3	647 (20.0%)	430 (17.0%)	217 (30.3%)	
Group4	648 (20.0%)	552 (21.9%)	96 (13.4%)	
Group5	648 (20.0%)	541 (21.4%)	107 (14.9%)	

Table 1: Descriptive statistics by MADRS score groups (categorical variables are presented using the number (percentage); numerical variables are presented using the average (standard deviation). Difference comparison between two level MADRS SCORE (< 40 or ≥ 40) are conducted using t-test or ANOVA

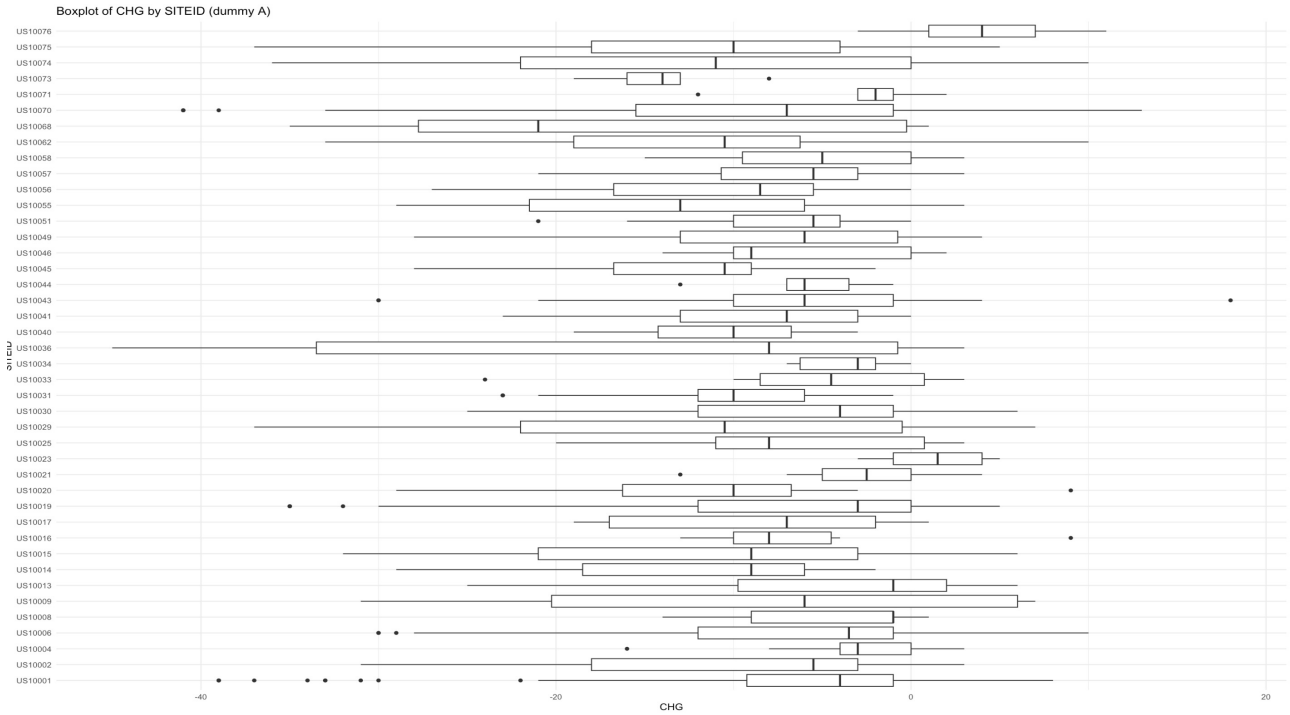

Fig. 2: MADRS Score in The Treatment of Placebo in Different Sites

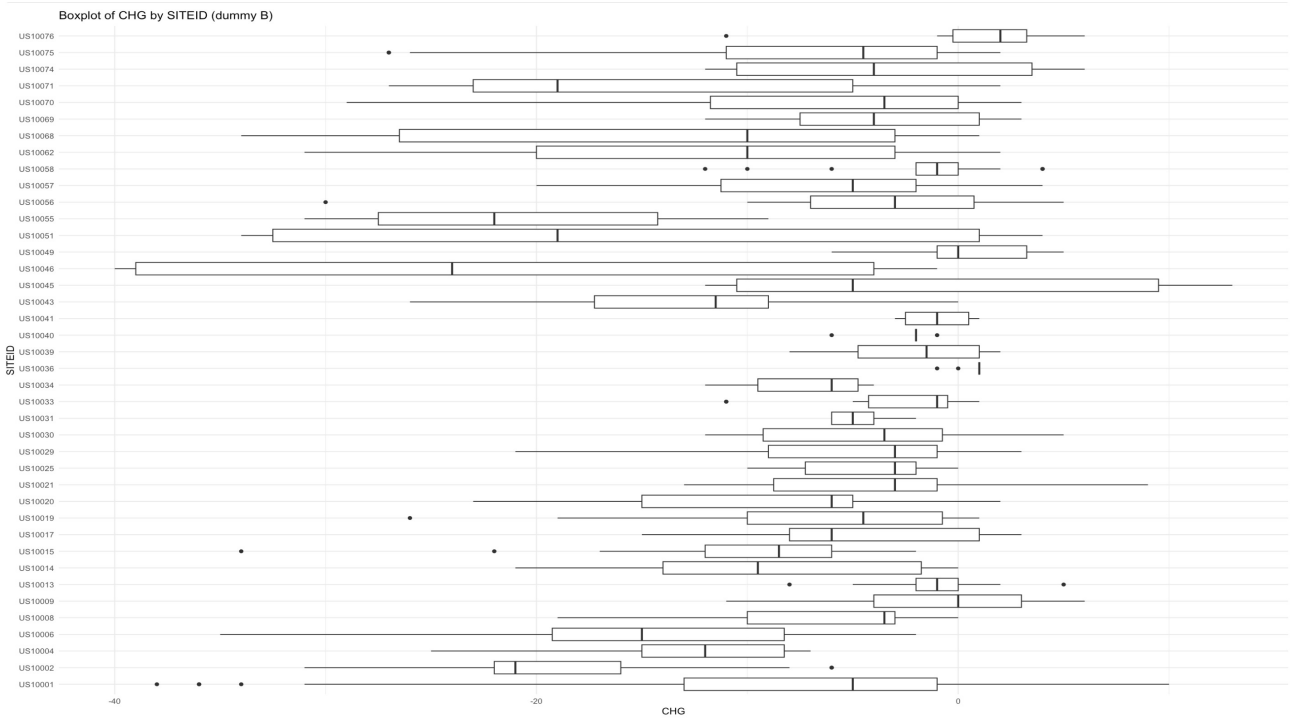

Fig. 3: MADRS Score in The Treatment of Esketamine (56 mg) in Different Sites

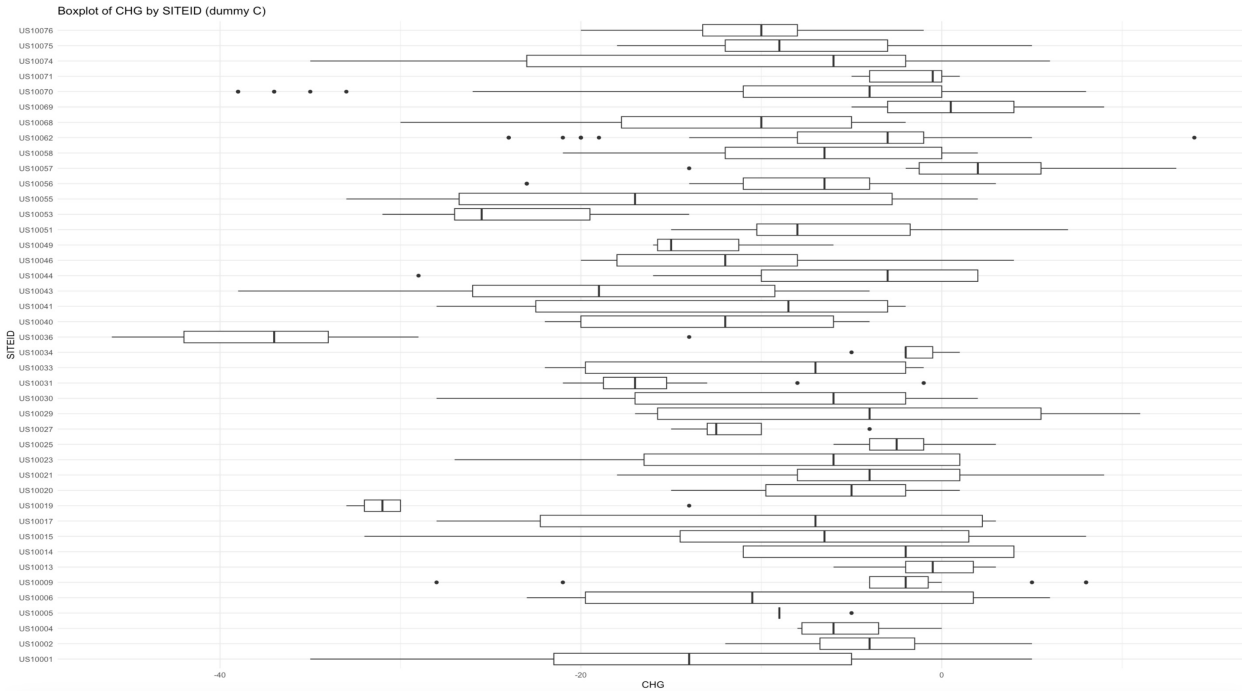

Fig. 4: MADRS Score in The Treatment of Esketamine (84 mg) in Different Sites

4.2 Model Results

4.2.1 MMRM Result

The results presented are derived from a MMRM that examines depression scores (AVAL) across multiple time points (AVISITN). The model includes the following factors: treatment effects (TRT01A, with three treatment options: placebo, Esketamine 56 mg, and Esketamine 84 mg), site effects (SITEID), demographic factors (RACE and AGE), baseline depression severity (BaseSituation), and an unstructured covariance matrix to account for within-subject correlations across visits for patients with depression.

Model Formula:

AVAL ~ AVISITN:TRT01A + TRT01A + SITEID + RACE + AGE + BaseSituation + us(AVISITN | USUBJID)

The analysis included 3,214 observations from 476 subjects across a maximum of 7 time points (from baseline to endpoint). The model utilized the Satterthwaite method for degrees of freedom and REML for inference, with asymptotic variance-covariance estimation.

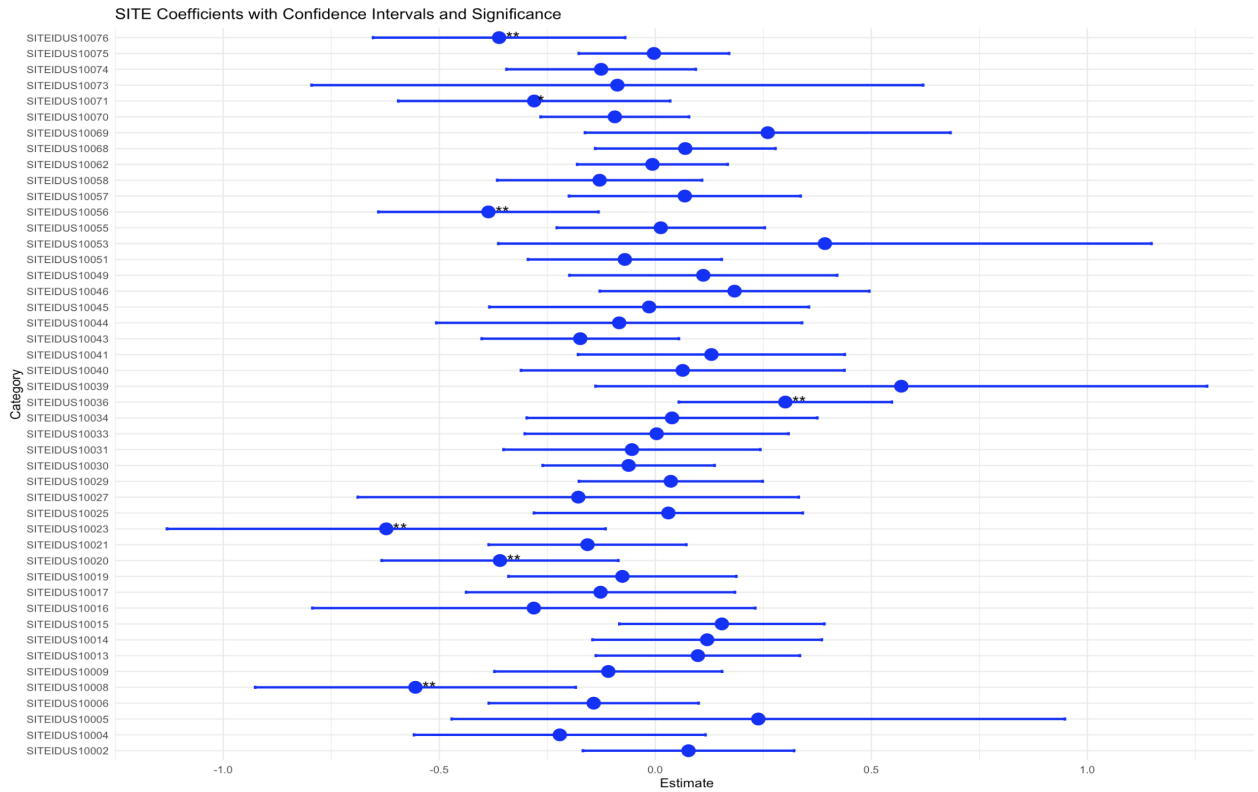

Fig. 5: 95 % Confidence Interval for the coefficients of Different Sites

Variable	CI	p_value
(Intercept)	1.806 (1.07, 2.542)	<0.001
TRT01ADummy B	0.002 (-0.079, 0.082)	0.967
TRT01ADummy C	0.048 (-0.034, 0.13)	0.251
SITEIDUS10002	0.077 (-0.168, 0.321)	0.538
SITEIDUS10004	-0.221 (-0.558, 0.116)	0.2
SITEIDUS10005	0.238 (-0.471, 0.948)	0.511
SITEIDUS10006	-0.143 (-0.385, 0.1)	0.251
SITEIDUS10008	-0.555 (-0.926, -0.184)	0.004
SITEIDUS10009	-0.109 (-0.372, 0.155)	0.419
SITEIDUS10013	0.099 (-0.138, 0.335)	0.414
SITEIDUS10014	0.12 (-0.145, 0.385)	0.376
SITEIDUS10015	0.154 (-0.083, 0.391)	0.204
SITEIDUS10016	-0.281 (-0.794, 0.231)	0.283
SITEIDUS10017	-0.127 (-0.438, 0.184)	0.425
SITEIDUS10019	-0.076 (-0.34, 0.188)	0.572
SITEIDUS10020	-0.359 (-0.633, -0.086)	0.01
SITEIDUS10021	-0.157 (-0.386, 0.072)	0.18
SITEIDUS10023	-0.623 (-1.13, -0.115)	0.017
SITEIDUS10025	0.03 (-0.281, 0.341)	0.849
SITEIDUS10027	-0.178 (-0.688, 0.332)	0.494
SITEIDUS10029	0.036 (-0.177, 0.249)	0.741
SITEIDUS10030	-0.062 (-0.261, 0.137)	0.544
SITEIDUS10031	-0.054 (-0.351, 0.243)	0.722
SITEIDUS10033	0.003 (-0.302, 0.308)	0.984
SITEIDUS10034	0.039 (-0.297, 0.375)	0.821
SITEIDUS10036	0.301 (0.055, 0.548)	0.017
SITEIDUS10039	0.569 (-0.138, 1.277)	0.115
SITEIDUS10040	0.063 (-0.311, 0.438)	0.74
SITEIDUS10041	0.13 (-0.179, 0.439)	0.41
SITEIDUS10043	-0.174 (-0.402, 0.055)	0.137
SITEIDUS10044	-0.084 (-0.507, 0.34)	0.699
SITEIDUS10045	-0.014 (-0.384, 0.356)	0.94
SITEIDUS10046	0.183 (-0.129, 0.495)	0.25
SITEIDUS10049	0.111 (-0.199, 0.421)	0.482
SITEIDUS10051	-0.07 (-0.295, 0.154)	0.538
SITEIDUS10053	0.393 (-0.363, 1.148)	0.309
SITEIDUS10055	0.013 (-0.228, 0.253)	0.917
SITEIDUS10056	-0.386 (-0.641, -0.131)	0.003
SITEIDUS10057	0.068 (-0.2, 0.336)	0.617
SITEIDUS10058	-0.129 (-0.366, 0.108)	0.287
SITEIDUS10062	-0.007 (-0.181, 0.168)	0.941
SITEIDUS10068	0.069 (-0.139, 0.278)	0.514
SITEIDUS10069	0.26 (-0.163, 0.683)	0.228
SITEIDUS10070	-0.094 (-0.265, 0.078)	0.286
SITEIDUS10071	-0.28 (-0.595, 0.034)	0.082
SITEIDUS10073	-0.088 (-0.796, 0.62)	0.808
SITEIDUS10074	-0.125 (-0.344, 0.094)	0.263
SITEIDUS10075	-0.003 (-0.177, 0.171)	0.972
SITEIDUS10076	-0.361 (-0.653, -0.07)	0.016
RACEASIAN	-0.284 (-1.031, 0.464)	0.458
RACEBLACK OR AFRICAN AMERICAN	-0.567 (-1.299, 0.165)	0.13
RACEMULTIPLE	-0.648 (-1.418, 0.122)	0.1
RACENATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER	-0.437 (-1.319, 0.445)	0.332
RACENOT REPORTED	-0.271 (-1.084, 0.542)	0.514
RACEUNKNOWN	-0.484 (-1.506, 0.538)	0.353
RACEWHITE	-0.463 (-1.184, 0.258)	0.209
AGE	0.022 (-0.014, 0.057)	0.233

Table 2: Coefficients with 95% Confidence Intervals and p-values

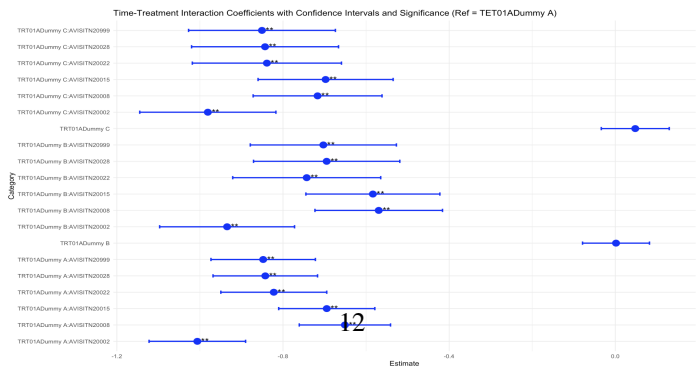

Fig. 6: 95 % Confidence Interval for the coefficients of Time-Treatment Interaction

Variable	CI	p_value
base_situation10-19	-2.34 (-2.708, -1.972)	<0.001
base_situation20-39	-0.857 (-0.937, -0.777)	<0.001
TRT01ADummy A:AVISITN20002	-1.006 (-1.122, -0.89)	<0.001
TRT01ADummy B:AVISITN20002	-0.935 (-1.097, -0.772)	<0.001
TRT01ADummy C:AVISITN20002	-0.981 (-1.145, -0.817)	<0.001
TRT01ADummy A:AVISITN20008	-0.651 (-0.761, -0.541)	<0.001
TRT01ADummy B:AVISITN20008	-0.57 (-0.723, -0.416)	<0.001
TRT01ADummy C:AVISITN20008	-0.717 (-0.872, -0.562)	<0.001
TRT01ADummy A:AVISITN20015	-0.695 (-0.81, -0.579)	<0.001
TRT01ADummy B:AVISITN20015	-0.584 (-0.745, -0.422)	<0.001
TRT01ADummy C:AVISITN20015	-0.697 (-0.86, -0.535)	<0.001
TRT01ADummy A:AVISITN20022	-0.822 (-0.95, -0.695)	<0.001
TRT01ADummy B:AVISITN20022	-0.743 (-0.921, -0.565)	<0.001
TRT01ADummy C:AVISITN20022	-0.839 (-1.018, -0.659)	<0.001
TRT01ADummy A:AVISITN20028	-0.843 (-0.968, -0.717)	<0.001
TRT01ADummy B:AVISITN20028	-0.695 (-0.871, -0.519)	<0.001
TRT01ADummy C:AVISITN20028	-0.843 (-1.02, -0.666)	<0.001
TRT01ADummy A:AVISITN20999	-0.848 (-0.973, -0.722)	<0.001
TRT01ADummy B:AVISITN20999	-0.703 (-0.879, -0.527)	<0.001
TRT01ADummy C:AVISITN20999	-0.851 (-1.027, -0.674)	<0.001

Table 3: Continued Table 1: Coefficients with 95% Confidence Intervals and p-values

Table 4: Unstructured Covariance Estimates

	0	2	8	15	22	28	999
0	0.1287	0.0874	0.0944	0.0802	0.0866	0.0829	0.0830
2	0.0874	0.8628	0.5757	0.5341	0.5563	0.5582	0.5632
8	0.0944	0.5757	0.7903	0.6412	0.6584	0.6251	0.6278
15	0.0802	0.5341	0.6412	0.8388	0.7806	0.7324	0.7313
22	0.0866	0.5563	0.6584	0.7806	1.0243	0.8904	0.8913
28	0.0829	0.5582	0.6251	0.7324	0.8904	1.0050	1.0005
999	0.0830	0.5632	0.6278	0.7313	0.8913	1.0005	1.0039

The model assessed three treatments (placebo as reference, 56 mg Esketamine, and 84 mg Esketamine). The main effect parameters for treatments B and C (0.002 and 0.048, respectively) were not statistically significant (p=0.967 and p=0.251, respectively), suggesting no overall difference in depression scores between treatments when controlling for other factors. However, the treatment-by-visit interactions revealed important temporal patterns in treatment effects since all treatments showed significant improvement over time (negative coefficients for visit interactions) as indicated by figure 6. For example, at the second visit day, all treatments showed substantial and significant reductions in depression scores (all p<0.001). Besides, treatment B appeared to have slightly smaller reductions compared to treatments A and C across most visits. By visits 20028 and 20999 (final assessment), the depression score reductions were -0.84 points (p<0.001), -0.70 points (p<0.001) and -0.85 points (p<0.001) respectively. This pattern suggests that treatments A and C may have slightly more sustained effects than treatment B, though all treatments demonstrated substantial and statistically significant improvement over time. Besides, different sites have shown different patterns in their confident intervals as shown on figure 5.

4.2.2 Emeans Evaluation

This study employed Mixed-Model Repeated Measures (MMRM) analysis to evaluate the efficacy of Esketamine nasal spray monotherapy in adults with treatment-resistant depression, with the primary endpoint being the change in MADRS total score from baseline to day 28 (table 5). Results demonstrated that both doses of Esketamine (56 mg and 84 mg) were significantly superior to placebo, with efficacy that was both statistically significant and clinically meaningful.

Table 5: Depression Scores by Treatment Group at Visit 28

TRT01A	emmean	SE	df	lower.CL	upper.CL
Dummy A	25.0	1.43	800	22.2	27.8
Dummy B	15.6	1.75	1000	12.2	19.0
Dummy C	16.1	1.63	981	12.9	19.3

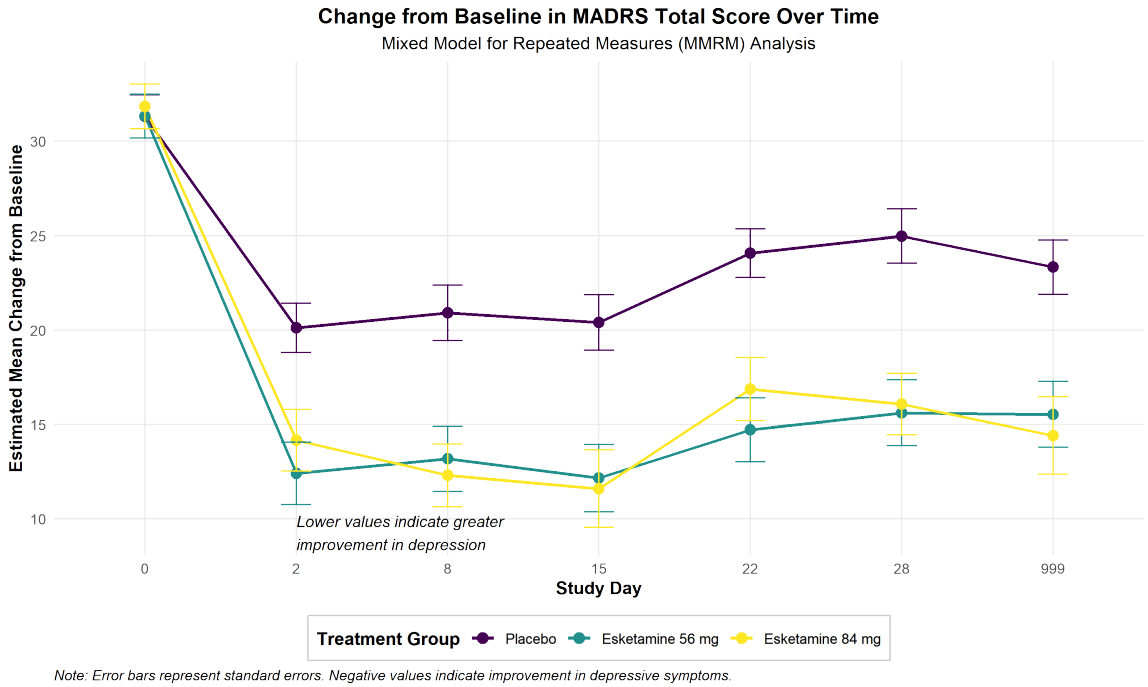

Fig. 7: MADRS Change Over Time

The study assessed the impact of Esketamine nasal spray on MADRS total scores of treatment-resistant depression patients across different time points from baseline to endpoint (as shown in Figure 7). Results indicated that the Esketamine treatment groups (56 mg and 84 mg) exhibited rapid symptom improvement early in the treatment course, with efficacy sustained through the study endpoint, significantly outperforming the placebo group (Figure 7).

In the analysis, data showed no significant differences in baseline MADRS scores among the three groups (Dummy A: 31.31, Dummy B: 31.33, Dummy C: 31.84, $p > 0.05$), indicating good baseline balancing.

By day 2, the Esketamine groups showed significant score reductions, with the 56 mg group (Dummy B) reaching a mean of 12.41 (SE=1.66, 95% CI: 9.16-15.66) and the 84 mg group (Dummy C) reaching 14.16 (SE=1.63, 95% CI: 10.96-17.37), while the placebo group (Dummy A) only decreased to 20.12 (SE=1.30).

By day 8, efficacy was further consolidated, with means for Dummy B and C decreasing to 13.18 and 12.31, respectively, significantly lower than the placebo group's 20.92 ($p < 0.0001$), suggesting rapid drug efficacy within the first week.

Moreover, until the endpoint of the double-blind study period, both treatment groups maintained stable efficacy (Dummy B: 15.54, Dummy C: 14.41), while the placebo group rebounded to 23.34, further supporting the long-term effectiveness of Esketamine nasal spray.

Notably, the 95% CIs of the Esketamine groups did not overlap with the placebo group at any time point, indicating statistically significant efficacy.

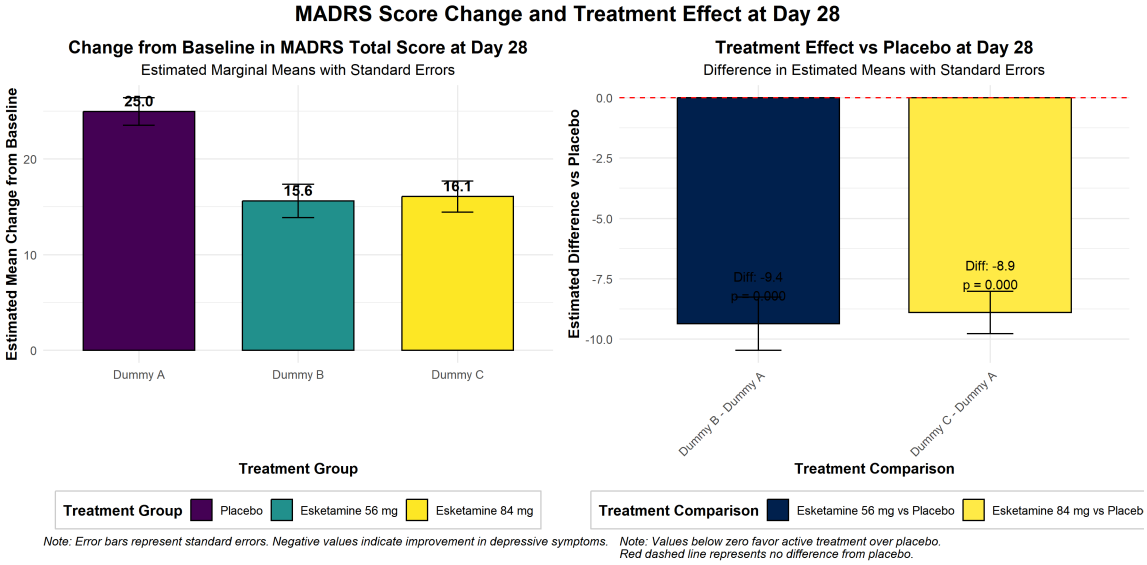

Fig. 8: MADRS change and Treatment Effect at Day 28

The placebo group (Dummy A) demonstrated an estimated marginal mean MADRS total score of 24.98 (SE=1.43, 95% CI: 22.16-27.79) at day 28, indicating limited symptom improvement.

The Esketamine 56 mg group (Dummy B) exhibited a mean score of 15.62 (SE=1.75, 95% CI: 12.19-19.05), representing a significant reduction from baseline.

The Esketamine 84 mg group (Dummy C) showed a mean score of 16.08 (SE=1.63, 95% CI: 12.88-19.28), with minimal difference compared to the 56 mg group, yet still significantly superior to placebo.

As illustrated in above Figure 8 (left panel), the confidence intervals for MADRS total scores across all treatment groups did not overlap with the placebo group, demonstrating the robustness of the therapeutic efficacy.

Figure 8 (right panel) presents data on treatment effect comparisons, wherein the estimated difference in least squares means between Esketamine 56 mg and placebo was -9.36 (SE=1.10, $p < 0.0001$), while the difference between Esketamine 84 mg and placebo was -8.90 (SE=0.88, $p < 0.0001$). These differences exceed the threshold of clinical significance for MADRS score reduction (typically established at 2 points). Furthermore, both comparisons maintained high statistical significance even after Dunnett's multiple comparison adjustment, substantiating the robust therapeutic efficacy of Esketamine.

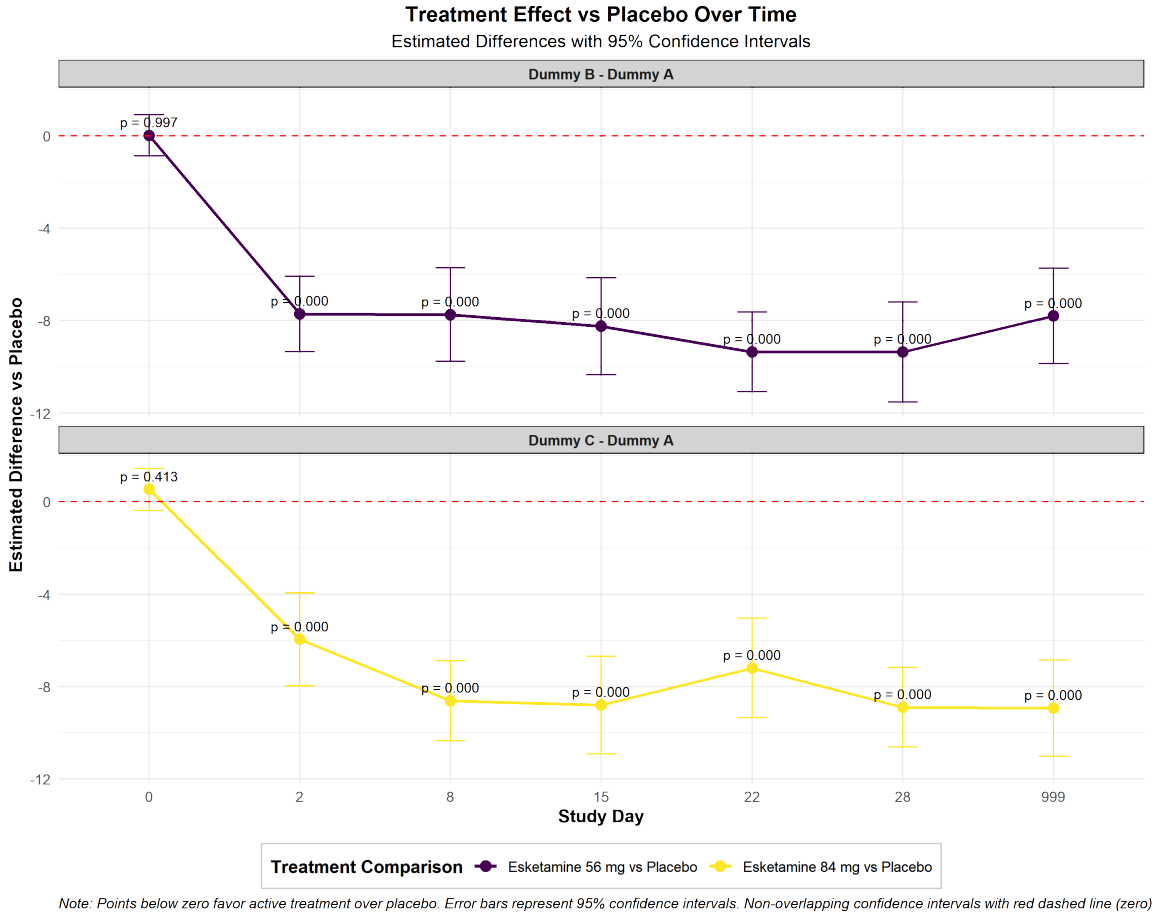

Fig. 9: Treatment effect forest plot

This study also utilized forest plots to visually demonstrate the treatment effect differences between Esketamine nasal spray (56 mg and 84 mg) and placebo across various time points. The analysis employed Mixed-Model Repeated Measures (MMRM), controlling for confounding factors such as study centers (SITEID) and baseline symptoms (basesituation), with p-values adjusted using Dunnett's multiple comparison method. Results indicated that both dosage groups significantly outperformed placebo at all assessment time points, with effect sizes exhibiting dynamic changes over time (Figure 9).

The forest plot revealed no significant differences in baseline scores between Esketamine groups and placebo (Dummy B vs A: p=0.997; Dummy C vs A: p=0.413), further confirming baseline equivalence. From day 2 through endpoint (AVISITN=2-999), all time point comparisons achieved high statistical significance ($p < 0.0001$). On day 2, Dummy B (56 mg) demonstrated an effect size of -7.71 (SE=0.83, $p < 0.0001$), while Dummy C (84 mg) showed an effect size of -5.96 (SE=1.03, $p < 0.0001$), indicating rapid early onset of action. By day 28, effect sizes remained stable, with Dummy B showing a difference of -9.36 (SE=1.10) and Dummy C showing -8.89 (SE=0.88), both significantly superior to placebo ($p < 0.0001$). At endpoint (formal cessation of the double-blind study phase), effect sizes maintained stability (Dummy B: -7.80, Dummy C: -8.93, $p < 0.0001$), supporting long-term efficacy.

Moreover, after applying Dunnett's correction for the two comparisons, all p-values maintained high statistical significance (adjusted $p < 0.0001$), reducing the risk of Type I error.

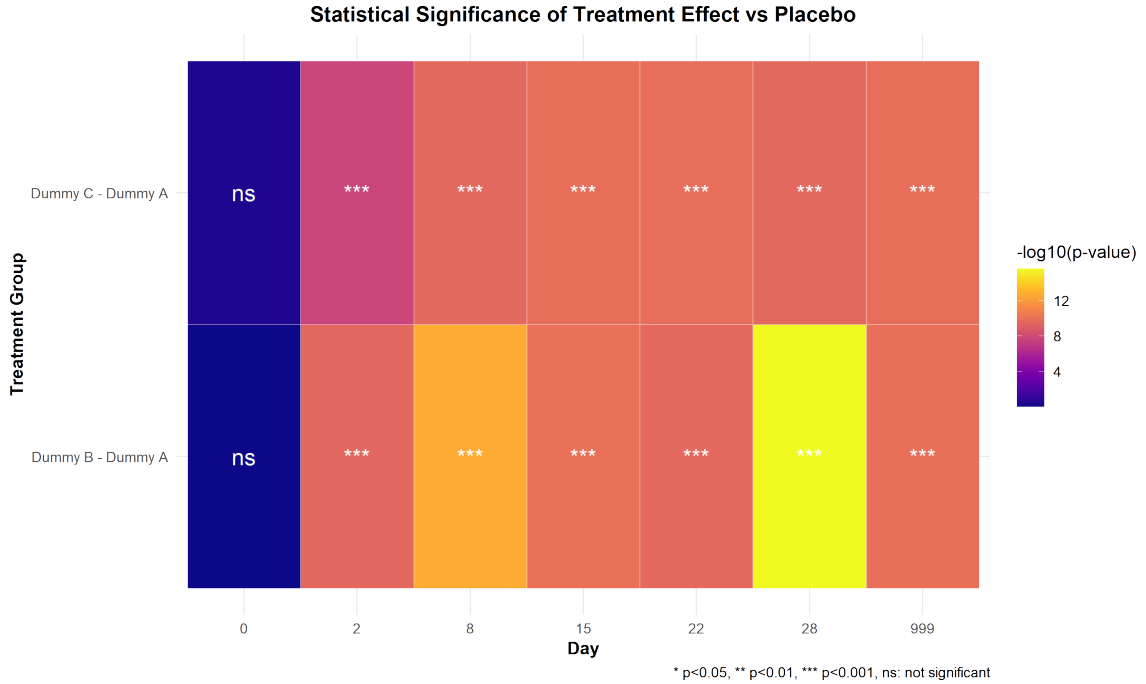

Fig. 10: Treatment significance heatmap

This study employed a heatmap to visually demonstrate the significance of treatment effects between Esketamine nasal spray (56 mg and 84 mg) and placebo across different time points. The analysis utilized Mixed-Model Repeated Measures (MMRM) methodology, and enhanced the visual differentiation of smaller p-values through negative logarithmic transformation ($-\log_{10}(p)$), thereby quantifying the statistical significance of Esketamine nasal spray efficacy. In the heatmap, darker regions correspond to higher -log(p) values (i.e., lower p-values), providing an intuitive representation of significance levels. Results depicted in the heatmap similarly demonstrated that both dosage groups exhibited highly significant differences at all assessment time points post-treatment (except baseline), with effect intensity dynamically changing over time (Figure 10).

4.3 Model's Performance Evaluation and Comparison

Model Evaluation:

This study employed diagnostic plots to verify the validity of statistical assumptions, ensuring the robustness of inferential conclusions. The diagnostic plots, presented in Figure 11, include residual-versus-fitted value plots and normal Q-Q plots, with the following specific analyses:

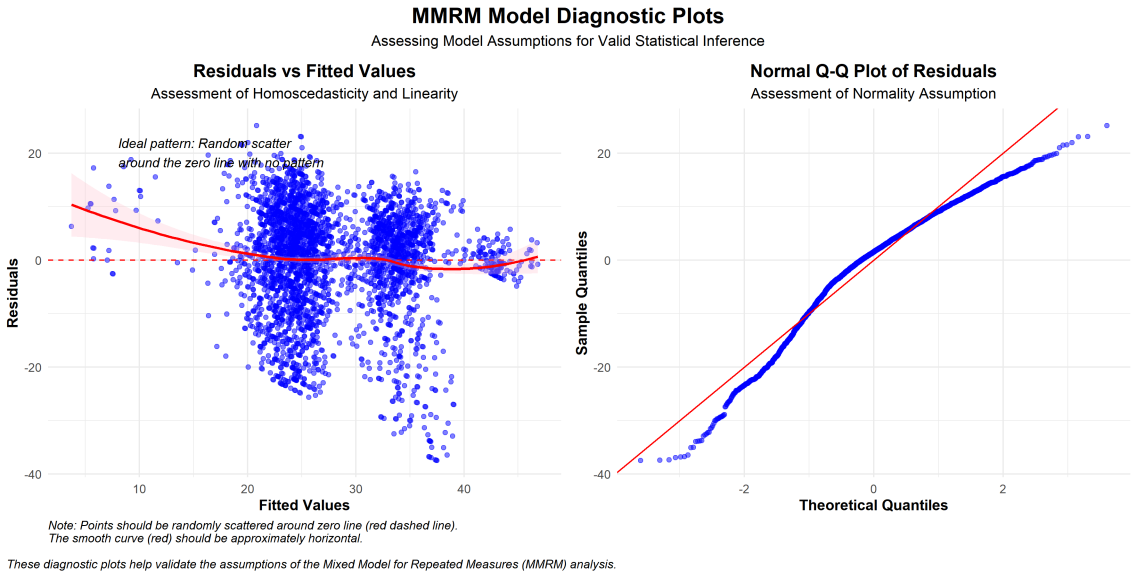

Fig. 11: Work Flow Chart

The residual-versus-fitted value plot (Figure 11, left panel) was utilized to examine homoscedasticity and linearity assumptions. This plot demonstrates that residuals are randomly dispersed around the zero line (red dashed line), without notable clustering or funnel-shaped trends, indicating that the data satisfy the homoscedasticity assumption. The red smoothing curve appears approximately horizontal, further supporting

the constancy of residual variance and absence of evident non-linear relationships. This indicates the absence of systematic associations between fitted values and residuals, suggesting that the model adequately captures the variation in the data without omitting important predictor variables or interaction effects.

The normal Q-Q plot (Figure 11, right panel) was employed to assess normality assumptions. The plot shows that sample quantiles are distributed primarily along the diagonal line, with only minor deviations in the tails, indicating that residuals approximately follow a normal distribution. These slight tail deviations may be attributable to extreme values or the relatively large sample size, and do not compromise the overall robustness of the inferences. Observation confirms that the overall pattern of the Q-Q plot conforms to the normality assumption, supporting the validity of subsequent p-values and confidence intervals.

The random distribution of residuals, constant variance, and approximate normality collectively validate the appropriateness of the MMRM model. Consequently, the statistical significance of the study results (e.g., between-treatment group differences with $p < 0.0001$) demonstrates high credibility.

4.4 Subgroup Analysis

This study employed the Mixed Model for Repeated Measures (MMRM) to evaluate the efficacy of Esketamine (56mg and 84mg) compared to a placebo in patients with different genders, different levels of depression (moderate and severe) and different age groups (18-24 years, 24-55 years, and ≥55 years). The primary outcome measure was the total score on the Montgomery-Åsberg Depression Rating Scale (MADRS), with lower scores indicating less severe depressive symptoms. The results showed that Esketamine demonstrated more significant improvements in depressive symptoms across all subgroups.

4.4.1 Basis for Grouping

We chose gender age and individual depression severity (moderate and severe) as the basis for subgroup analysis. This decision was influenced by the findings of Cipriani et al. (2018), which suggested that age gender and depression severity might affect the therapeutic response to antidepressants. For instance, the study noted differences in the response to antidepressants among patients of different gender, age groups and severity levels, which could reflect the heterogeneity of depression mechanisms and etiologies. Following the MADRS scoring criteria by Zimmerman et al. (2004), patients were categorized into moderate depression (AVAL < 40) and severe depression (AVAL > 60) groups to explore the differential effects of Esketamine on patients

with varying severity levels. This classification is clinically significant because patients with severe depression often exhibit greater treatment resistance and a higher risk of suicide. By comparing the therapeutic effects of Esketamine in these two groups, we can determine whether it can compensate for the current treatment deficiencies, especially for patients with treatment-resistant severe depression. Additionally, this analysis helps guide clinical decision-making by identifying which patient groups are most likely to benefit from Esketamine treatment.

4.4.2 Statistical Methods and Error Representation

In all the graphs of this study, each data point represents the estimated marginal means obtained from the Mixed Model for Repeated Measures (MMRM) analysis, with vertical error bars indicating the standard error (SE). The standard error reflects the precision and reliability of the estimated means; shorter error bars indicate more precise estimates. Overlapping error bars usually suggest that the differences between groups are statistically insignificant, but definitive statistical significance should be confirmed through formal hypothesis testing. When interpreting treatment effects, we not only considered the trend of point estimates but also integrated the uncertainty information provided by the error bars to ensure the statistical robustness of our conclusions.

4.4.3 Sex

After including the confounding variable sex in our model, the result did not show big difference as table 3 2, as indicated by **??**. Besides, the model selection criterias AIC (3856.3) and BIC (3972.9) are both higher than the model in the 4.2.1 (AIC = 3852.3 ; BIC =3969.0). This means the sex did not show great difference in the effectiveness of the drug.

Table 6: Sex difference result using MMRM

Variable	Coefficient	95% CI	p-value
(Intercept)	1.801	(1.065, 2.538)	< 0.001
TRT01ADummy B	0.002	(-0.079, 0.082)	0.96678
TRT01ADummy C	0.049	(-0.032, 0.131)	0.23741
SEXM	-0.035	(-0.105, 0.035)	0.33304
base_situation10-19	-2.334	(-2.702, -1.965)	< 0.001
base_situation20-39	-0.853	(-0.933, -0.772)	< 0.001
TRT01ADummy A:AVISITN2	-1.006	(-1.122, -0.89)	< 0.001
TRT01ADummy B:AVISITN2	-0.935	(-1.097, -0.772)	< 0.001
TRT01ADummy C:AVISITN2	-0.981	(-1.145, -0.817)	< 0.001
TRT01ADummy A:AVISITN8	-0.651	(-0.761, -0.541)	< 0.001
TRT01ADummy B:AVISITN8	-0.57	(-0.723, -0.416)	< 0.001
TRT01ADummy C:AVISITN8	-0.717	(-0.872, -0.561)	< 0.001
TRT01ADummy A:AVISITN15	-0.695	(-0.81, -0.579)	< 0.001
TRT01ADummy B:AVISITN15	-0.584	(-0.745, -0.422)	< 0.001
TRT01ADummy C:AVISITN15	-0.697	(-0.86, -0.535)	< 0.001
TRT01ADummy A:AVISITN22	-0.822	(-0.95, -0.695)	< 0.001
TRT01ADummy B:AVISITN22	-0.743	(-0.921, -0.565)	< 0.001
TRT01ADummy C:AVISITN22	-0.839	(-1.018, -0.659)	< 0.001
TRT01ADummy A:AVISITN28	-0.843	(-0.968, -0.717)	< 0.001
TRT01ADummy B:AVISITN28	-0.695	(-0.871, -0.519)	< 0.001
TRT01ADummy C:AVISITN28	-0.843	(-1.02, -0.666)	< 0.001
TRT01ADummy A:AVISITN999	-0.848	(-0.973, -0.722)	< 0.001
TRT01ADummy B:AVISITN999	-0.703	(-0.879, -0.527)	< 0.001
TRT01ADummy C:AVISITN999	-0.85	(-1.027, -0.674)	< 0.001

4.4.4 Depression Degree

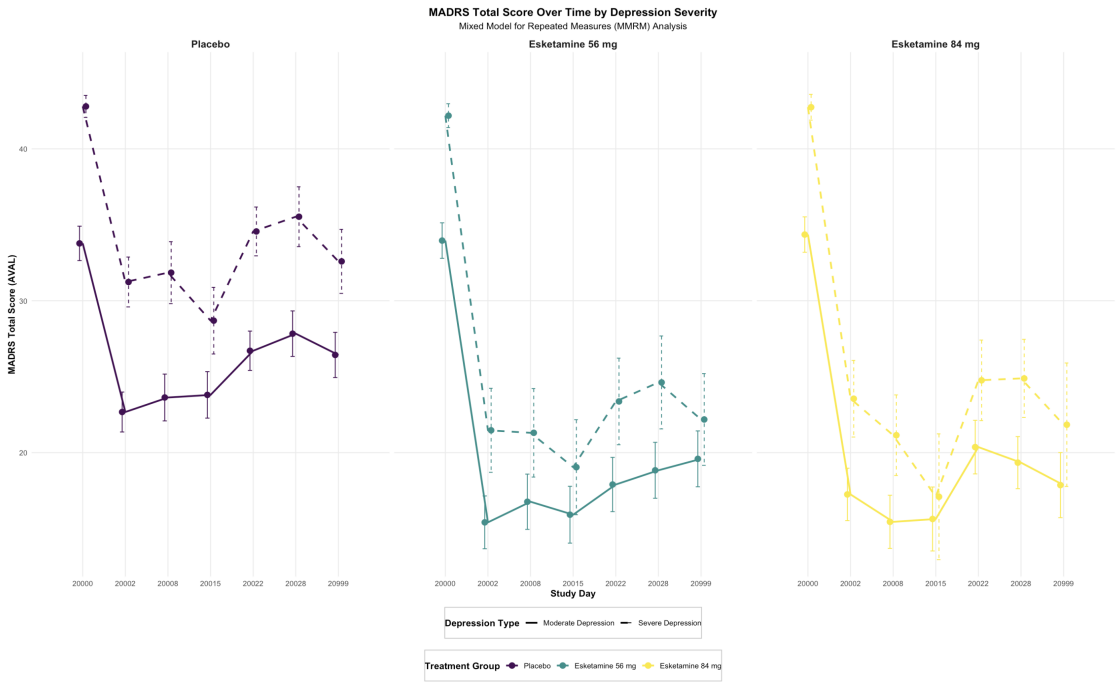

Fig. 12

Our study demonstrated that Esketamine showed significant therapeutic effects in patients with different levels of depression. In the placebo group, patients with moderate depression experienced a reduction in

MADRS scores from approximately 33 at baseline to about 22, while those with severe depression saw scores decrease from about 44 to around 28. Despite the larger improvement in severe depression patients, their final scores remained higher than those with moderate depression. In contrast, the Esketamine treatment group exhibited more pronounced therapeutic effects. Among patients receiving 56mg of Esketamine, those with moderate depression saw scores drop from approximately 15 to about 19, and those with severe depression from about 19 to around 15. Similarly, the 84mg dose group showed comparable results, with moderate depression patients experiencing a reduction from approximately 15 to about 19, and severe depression patients from about 21 to around 15. It is noteworthy that both doses of Esketamine led patients to achieve similar final scores, indicating that Esketamine may have greater clinical significance for patients with severe depression.

4.4.5 Age

Fig. 13

Age subgroup analysis revealed differential therapeutic effects of Esketamine across different age groups. In the placebo group, elderly patients (≥55 years) had the highest baseline MADRS scores (approximately 57), which remained relatively high after treatment (about 50); whereas the 18-24 year and 24-55 year age groups showed similar final scores of about 26. The Esketamine 56mg group demonstrated the most notable age-related differences: elderly patients experienced a rapid initial drop in scores from about 36 to about 2,

stabilizing around 15, indicating a potential special therapeutic effect for this group. The 84mg group showed consistent therapeutic responses across all age groups: both the 18-24 year and 24-55 year age groups saw scores drop to about 18, while the elderly group experienced a reduction from the highest baseline (about 51) to about 26. These results suggest that Esketamine, particularly the 56mg dose, may be an optimal treatment option for elderly patients with depression.

4.4.6 Summary

Our study concludes that Esketamine holds significant potential for treating depression, with notable differential effects across subgroups. Firstly, the inclusion of gender in the model did not significantly alter the therapeutic outcomes, suggesting that Esketamine's effectiveness is relatively consistent across genders. Secondly, patients with severe depression exhibited substantial improvement with Esketamine treatment, highlighting its value for severe or treatment-resistant cases. Lastly, age also emerged as a critical factor, with elderly patients showing a particularly strong initial response to the 56mg dose. These findings underscore the importance of considering these factors in optimizing treatment strategies.

5 Discussions

Dose-Response Relationship Analysis

Our study results indicate that Esketamine exhibits significant antidepressant effects at both 56mg and 84mg dosages, yet the relationship between dosage and effect is not perfectly linear. Specifically, most patient subgroups achieved notable clinical improvement at the 56mg dose, and increasing the dosage to 84mg did not yield a proportional increase in benefits. This finding holds significant implications for clinical practice, suggesting that treatment should prioritize starting at the lower effective dose, especially for certain populations who may be more sensitive to medication, such as the elderly. Additionally, we observed that patients with severe depression might benefit more from slightly higher doses, indicating that clinicians should adjust the dosage based on the severity of depression to balance optimal treatment efficacy and safety.

Clinical Significance of Depression Severity Differences

The significant response of patients with severe depression to Esketamine in this study is of considerable clinical importance. Traditionally, patients with severe depression have poorer treatment outcomes, lower remission rates, and longer times to remission. Esketamine's ability to rapidly and significantly improve

symptoms in this group could potentially shorten the duration of patient suffering and reduce the risk of suicide. Moreover, although the final MADRS scores of patients with severe depression were slightly higher than those with moderate depression, the relative improvement was greater, suggesting that Esketamine may specifically address neurotransmitter imbalances associated with severe depression by modulating the glutamatergic system. Therefore, clinicians should consider Esketamine as an early intervention option for patients with severe depression, particularly those who have not responded well to traditional treatments. Age-Related Differences in Treatment Response and Mechanism Exploration The unique response pattern of elderly patients to the 56mg dose of Esketamine is worthy of further investigation. Possible mechanisms include: (1) distinctive changes in the glutamatergic system of elderly patients, making them more sensitive to NMDA receptor antagonism; (2) elevated levels of chronic inflammation in the elderly, which Esketamine may counteract due to its potential anti-inflammatory effects; (3) the presence of neurodegenerative changes often accompanying depression in the elderly, where Esketamine's neuroprotective and neurogenic promoting effects may be more pronounced. Meanwhile, the consistent response of young and middle-aged patients to both doses also provides a reference for clinical medication use, indicating that the 56mg dose may be sufficient to achieve optimal efficacy and avoid unnecessary side effects associated with higher doses.

This study, through rigorous MMRM analysis, confirms the significant efficacy of Esketamine across different depression severity and age groups. Notably, the 56mg dose of Esketamine shows particularly significant effects in elderly patients aged 55 and above, offering a new treatment option for this traditionally challenging demographic. Its effectiveness in both moderate and severe depression patients demonstrates its broad applicability. These findings provide clinicians with important evidence-based support for adopting personalized treatment strategies based on patient characteristics. Future research should focus on Esketamine's long-term efficacy, maintenance treatment strategies, and combined applications with other antidepressant treatments to further optimize treatment plans. In summary, Esketamine, as an innovative treatment option, shows potential to improve clinical outcomes for various depression patients, offering new hope for the management of depression.

Difference of The Effectiveness Through Sites

- Limitations

This study has limitations, including a limited sample size and relatively short observation period. Future studies should expand the sample size and extend the follow-up duration to assess the long-term efficacy

and safety of Esketamine. Additionally, this study did not delve into the differential mechanisms of action of Esketamine; subsequent research should incorporate biomarkers and neuroimaging techniques to further elucidate its differentiated mechanisms of action across various populations.

6 Conclusion

This study demonstrated that Esketamine, at both 56 mg and 84 mg doses, provided rapid and sustained antidepressant effects in adults with treatment-resistant depression, with significant reductions in MADRS scores compared to placebo by day 28 and throughout the study endpoint. While both doses showed robust efficacy, the nonlinear dose-response relationship suggested that 56 mg may serve as an optimal starting dose for most patients, balancing efficacy and tolerability. Subgroup analyses across sex, depression severity and age subgroups confirmed the consistency of treatment effects, supported by model diagnostics validating the robustness of the MMRM framework, including residual plots confirming homoscedasticity and normality assumptions. The findings underscore esketamine's potential as a rapid-acting intervention for diverse populations, though longer-term studies are needed to evaluate sustained benefits and safety. These results highlight the importance of tailored dosing strategies and reinforce the clinical relevance of esketamine in addressing unmet needs in treatment-resistant depression.

7 Reference

Cipriani, A., Furukawa, T. A., Salanti, G., et al. (2018). Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. The Lancet, 391(10128), 1357-1366.

Laird, N. M., Ware, J. H. (1982). Random-effects models for longitudinal data. Biometrics, 38(4), 963–974.

Sabanes Bove, D., Li, L., Dedic, J., Kelkhoff, D., Kunzmann, K., Lang, B., Stock, C., Wang, Y., James, D., Sidi, J., Leibovitz, D., Sjoberg, D. (2025). mmrm: Mixed Models for Repeated Measures. R package version 0.3.14.9002. Available at: https://openpharma.github.io/mmrm/ (Accessed: 7 April 2025).

Zimmerman, M., Chelminski, I., Posternak, M. (2004). A review of studies of the Montgomery–Asberg Depression Rating Scale in controls: implications for the definition of remission in treatment studies of depression. International Clinical Psychopharmacology, 19(1), 1–7.

Appendix A

Codes are provided here:

https://hzzzzz411.github.io/MMRM-Analysis/

Acknowledgements. We extend our heartfelt thanks to Dr. Mu He for his invaluable guidance and for facilitating our communication with the Tigermed team. His support has been instrumental in providing us with the opportunity to learn and engage in meaningful discussions. Additionally, we are deeply appreciative of the consistent and well-processed data provided by Tigermed, which has been crucial for our analysis. Besides, thanks to our hard work and enthusiasm about biomedical statistics.