



# Evaluation of Individual Items of the Patient Health Questionnaire (PHQ-9) and Montgomery-Asberg Depression Rating Scale (MADRS) in Adults with Treatment-Resistant Depression Treated with Esketamine Nasal Spray Combined with a New Oral Antidepressant

Lysbeth Floden<sup>1</sup> · Stacie Hudgens<sup>1</sup> · Carol Jamieson<sup>2</sup> · Vanina Popova<sup>3</sup> · Wayne C. Drevets<sup>4</sup> · Kimberly Cooper<sup>5</sup> · Jaskaran Singh<sup>4,6</sup>

Accepted: 14 March 2022 / Published online: 20 April 2022  
© The Author(s) 2022

## Abstract

**Objective** The objective of this study was to determine which symptoms measured by the Patient Health Questionnaire (PHQ-9) and Montgomery-Asberg Depression Rating Scale (MADRS) improve in those treated with esketamine nasal spray in combination with oral antidepressant (AD) compared with those treated with placebo plus AD for adult patients with treatment-resistant depression (TRD). These results complement the interpretation of PHQ-9 and MADRS total scores.

**Methods** The TRANSFORM 2 study evaluated the efficacy and safety of esketamine nasal spray in combination with AD. This post-hoc analysis used PHQ-9 and MADRS data to evaluate symptom changes. The total scores change and proportions of individual item change scores on the PHQ-9 and MADRS were evaluated at days 15 and 28; analysis of variance was used to test differences on total scores. Generalized estimation equations of logistic regression models were used to estimate the likelihood of improvement on instrument items.

**Results** The mean total score reduction of the PHQ-9, indicating improvement, was greater in the esketamine plus AD arm compared with placebo plus AD at day 15 ( $-1.8$ ;  $p = 0.045$ ) and day 28 ( $-2.8$ ;  $p = 0.006$ ). Proportions of those who improved ( $\geq 1$  point on a 4-point scale and  $\geq 2$  points on a 7-point scale for the PHQ-9 and MADRS, respectively) was greater in the esketamine plus AD group compared with the placebo plus AD group across all items. The odds of improving for those in the esketamine plus AD group compared with the placebo plus AD group were over two times greater on the PHQ-9 items: “Little interest/pleasure in things” (OR 2.252, 95% CI 1.165–4.355); “Feeling down, depressed, or hopeless” (OR 2.767, 95% CI 1.400–5.470); and “Feeling tired or having little energy” (OR 2.171, 95% CI 1.153–4.087). The mean reduction in total scores on the MADRS, indicating improvement, was numerically greater at day 15 ( $-2.0$ ;  $p = 0.189$ ) and statistically significantly greater at day 28 ( $-4.4$ ;  $p = 0.017$ ) in the esketamine plus AD arm compared with placebo plus AD. The odds of improving for those in the esketamine plus AD group compared with the placebo plus AD group were over two times greater on the MADRS items measuring “Apparent sadness” (OR 2.007, 95% CI 1.096–3.674); and “Inability to feel” (OR 2.099, 95% CI 1.180–3.735).

**Conclusion** Improvement in mean total scores in those treated with esketamine plus AD compared with placebo plus AD are important results to confirm efficacy. The odds of improving in those treated with esketamine plus AD was at least two times greater than with placebo plus AD on three patient- and two clinician-reported individual symptoms of TRD. These findings provide patient-relevant quantification of the esketamine plus AD treatment benefit, adding understanding as to which symptoms are most improved with treatment.

**Trial Registration** ClinicalTrials.gov identifier: NCT02418585; first posted 16 April 2015.

## Key Points

Evaluating group-level total score change on the Patient Health Questionnaire (PHQ-9) and Montgomery-Asberg Depression Rating Scale (MADRS) are important assessments of efficacy in adult patients with TRD treated with esketamine plus AD compared with placebo plus AD. Additionally, assessing the individual items measuring symptoms aids interpretation of treatment benefits.

Adult patients treated with esketamine plus AD compared with placebo plus AD had improved total scores on both the PHQ-9 and MADRS, and were over two times more likely to improve on symptoms such as having little interest/pleasure in things; feeling down, depressed, or hopeless; and feeling tired or having little energy, as measured by the PHQ-9. Similarly, adult patients treated with esketamine plus AD compared with placebo plus AD were over two times more likely to improve on symptoms such as apparent sadness and inability to feel, as measured by the MADRS.

## 1 Introduction

Major depressive disorder (MDD) or depression is a leading cause of global long-term disability and is among the top ten reasons for years lived with disability in most countries [1]. Among those who receive treatment for depression, about 30% do not respond to at least two different pharmacological treatments and are diagnosed with treatment-resistant depression (TRD) [2]. This diagnosis contributes to the disease burden, as adult patients with TRD manifest lower health-related quality of life, increased mortality and morbidity, higher relapse rates within a year of remission, and increased treatment costs relative to adult patients with MDD that is more treatment responsive [3, 4].

Prior to the 2019 FDA approval of esketamine nasal spray for treatment of TRD and subsequent approval by numerous health authorities worldwide, individuals diagnosed with TRD had limited treatment options, with the fluoxetine/olanzapine combination being the only treatment approved in the US for TRD management [4]. Interventions including combining existing pharmacological treatment with non-pharmacological, psychological, or non-psychological methods are mostly ineffective or associated with significant tolerability or safety issues [5].

Esketamine, the S-enantiomer of ketamine racemate and an *N*-methyl-D-aspartate receptor antagonist, has been

recently approved as a nasal spray for the treatment of TRD [6–9]. The TRANSFORM 2 study evaluated the efficacy, safety, and tolerability of esketamine nasal spray (56 mg and 84 mg) used in combination with a newly prescribed open-label (OL) oral antidepressant (AD) in patients with TRD. Esketamine nasal spray was approved in the US in early 2019 as a Schedule III controlled substance to be administered under the direct supervision of a healthcare provider for treatment of TRD and was subsequently approved in other countries.

Patient-reported outcome (PRO) instruments are used for the detection of depression, assessment of severity and disease burden, and guidance of treatment selection in a clinical setting. The 9-item Patient Health Questionnaire (PHQ-9) is a PRO instrument that is used to measure MDD symptoms, as described by the patient, and measure efficacy over time [10]. This 9-item questionnaire is an efficient tool for evaluating depression severity and consists of the nine criteria upon which the diagnosis of depressive disorders is based (Diagnostic and Statistical Manual of Mental Disorders, 5th edition [DSM-5]) [11]. The Montgomery-Asberg Depression Rating Scale (MADRS), a 10-item questionnaire reported by the clinician, also mirrors the DSM-5 criteria.

The item scores on these clinical outcome assessment (COA) instruments are summed together for a total score measuring depression severity. These COA scales are composite measures comprising multiple items that assess depression, but items may contribute differently to overall depression severity. Clinicians, patients, and researchers often attempt to interpret the total scores without understanding the contribution of the individual items [12], which in the PHQ-9 and MADRS are also symptoms. Treatment effects can vary across symptoms and may differ depending on the mechanism of action of the intervention. The changes observed in a total score may result from changes in only a few symptoms but are generally interpreted as a change in the entire construct, symptoms of depression. Moreover, items measuring specific symptoms may have varying degrees of clinical importance, making evaluation of single items relevant to interpretation of the total score. For example, on the PHQ-9, item 9 reads, “Thoughts you would be better off dead, or of hurting yourself.” This reflects a more severe symptom of depression than item 7, which reads, “Trouble concentrating on things.” Nevertheless, both items are equal contributors to the total score according to the PHQ-9 scoring algorithm. The measurement of PRO symptom experiences via single items on these instruments help to interpret and compare the efficacy of treatment strategies.

Results presented are from prespecified, post-hoc analyses of data from the phase III TRANSFORM 2 randomized clinical trial assessing the efficacy of esketamine nasal spray for TRD. The TRANSFORM 2 study was chosen from other trials assessing esketamine efficacy to keep the demographic

composition similar to those most affected by TRD; because the study design allowed flexible dosing, which aligns with real-world use; and because the analysis of the primary endpoint was positive. The analyses evaluate patient-reported and clinician-reported improvements in depressive symptoms among adult patients with TRD receiving esketamine nasal spray plus a newly initiated oral AD (esketamine plus AD) versus those who received a newly initiated oral AD (active comparator) plus intranasal placebo (placebo plus AD).

The objective of the analyses was to determine whether individual items in the PHQ-9 and MADRS instruments that measure symptoms show differences by treatment arm over the course of treatment for patients with TRD. The study was registered at ClinicalTrials.gov (NCT02418585).

## 2 Methods

### 2.1 Study Characteristics

TRANSFORM-2 was a phase III, double-blind, multicenter, active-controlled study to evaluate the efficacy, safety, and tolerability of flexible doses of esketamine nasal spray (56 mg or 84 mg). The study was approved by the local ethics committee, and written informed consent was obtained. The Institutional Review Board (IRB) reviewed the study protocol, and all adverse events were reported to the IRB in compliance with the sponsor's standard operating procedures.

All patients included in the study met the DSM 5 diagnostic criteria for MDD without psychotic features. This diagnosis was determined by clinical assessment and confirmed by the Mini-International Neuropsychiatric Interview. As part of the inclusion criteria, patients had an Inventory of Depressive Symptomatology—Clinician-rated 30-item total score of  $\geq 34$ , which corresponds to moderate to severe MDD. These patients were treatment-resistant to at least two different antidepressant treatments within the current depressive episode. Inclusion criteria included the ability to read and understand study instruments and instruction. Additional details of the entrance criteria are described elsewhere [8].

Patients were randomized (1:1) to receive a double-blind flexible dose of esketamine nasal spray (56 mg or 84 mg) plus oral AD or oral AD plus intranasal placebo [8]. Intranasal treatment sessions (esketamine or placebo) occurred twice weekly for a 4-week period. Concurrently, patients initiated a new open-label oral antidepressant taken for the duration of the study.

### 2.2 Study Instruments

The PHQ-9 is a PRO instrument used to assess depression symptoms with the following nine items: Item 1—Little interest/pleasure in things; Item 2—Feeling down, depressed, or hopeless; Item 3—Trouble falling or staying asleep or sleeping too much; Item 4—Feeling tired or little energy; Item 5—Poor appetite or overeating; Item 6—Feeling bad about yourself; Item 7—Trouble concentrating on things; Item 8—Moving slowly or fidgety/restless; and Item 9—Thoughts you would be better off dead. This instrument can be used both as a screening tool and to evaluate the response to treatment for depression [10]. Each of the nine items are rated on a scale of 0 to 3 (0 = not at all, 1 = several days, 2 = more than half the days, and 3 = nearly every day) and the item responses are totaled to arrive at a final score (range of 0–27). Higher scores indicate a greater severity of depression. The recall period was 2 weeks. Prior to the initiation of the trial, validation work was undertaken to ensure the integrity of the instrument, as well as extensive feasibility testing of the electronic format. Electronic data collection devices were used to collect the PHQ-9 and other PRO data during onsite assessments and were presented to patients by trained study staff according to the schedule of events in the protocol. This was a multinational study, and the PROs were translated and administered in the language that was most appropriate for the patient's everyday language. The PROs also had to be conducted prior to other assessments for the study visit. Linguistic translations were conducted to The Professional Society for Health Economics and Outcomes Research guidelines [13].

The MADRS is a clinician-reported outcome (ClinRO) scale designed to measure depression severity and is responsive to changes due to AD treatment. The scale consists of ten items, each of which is scored from 0 (item not present or normal) to 6 (severe or continuous presence of the symptoms), for a total possible score of 60. Higher scores represent a more severe condition. Items are as follows: Item 1—Reported sadness; Item 2—Apparent sadness; Item 3—Inner tension; Item 4—Reduced sleep; Item 5—Reduced appetite; Item 6—Concentration difficulties; Item 7—Lassitude; Item 8—Inability to feel; Item 9—Pessimistic thoughts; and Item 10—Suicidal thoughts. MADRS data were collected by independent, remote raters < 2 days prior to the study visit. All raters underwent rater training for consistency and accuracy prior to trial initiation.

### 2.3 Statistical Analysis

The analytic population was defined as adult patients with COA assessments at any time point within the intent-to-treat population. These exploratory post-hoc analyses were performed using the statistical analysis system (SAS) Version 9.4.

Distributions and descriptive statistics of demographic and clinical characteristics were described at baseline for each treatment group and overall. Total scores and change from baseline on the PHQ-9 and MADRS were displayed by treatment arm from baseline to days 15 and 28; treatment differences were calculated using analysis of variance (ANOVA). Categorical distributions of change from baseline in each item of the PHQ-9 and MADRS were displayed graphically for days 15 and 28.

The definition of within-patient item-level improvement corresponded to the categorical shift of at least the magnitude of the total score meaningful change thresholds (MCTs). The within-patient MCTs for the PHQ-9 and MADRS are 6 points, or 22% of the total score range and 10 points, or 17% of the total score range, respectively [14]. For example, the minimal important difference (MID) of the MADRS is generally agreed to be 2 points [15]. Individual patients were classified as improved if they had a decrease of at least 1 point on the PHQ-9 items (representing a 25% shift, measured on a 4-point scale) or 2 points on the MADRS items (representing a 29% shift, measured on a 7-point scale). It is important to note that the MCT

differs from the threshold for clinical relevance both conceptually and often numerically. While the threshold for within-person meaningful change refers to the amount of change each individual needs to achieve to be classified as improving or deteriorating, the MID is a measure of clinical relevance used to judge clinical significance of mean difference in change between groups. The proportions of patients with categorical improvement from baseline of PHQ-9 and MADRS items, by treatment arm, were calculated at Days 15 and 28. Generalized estimation equations (GEE) of logistic regression models were used to estimate the likelihood of improvement. The GEE model of an 'improved' binary event was regressed on fixed effects of treatment arm, time point, and their interaction terms, as well as an R-side random effect with autoregressive correlation structure to account for repeated measures. Models generated odds ratios (OR) with 95% confidence intervals (CI) to compare the likelihood of 'improvement' between treatment groups.

### 3 Results

On average, patients were 46 years old (SD 11.89), non-Hispanic (93%), White (93%), and female (62%) with an average of 12 years (SD 10.2) since diagnosis of TRD. Patient demographic and clinical characteristics, including baseline PHQ-9 and MADRS scores, were similar across all treatment groups (Tables 1 and 2).

**Table 1** Baseline demographic and clinical information

Variable	Esketamine plus AD <i>n</i> = 114	Placebo plus AD <i>n</i> = 109	Overall <i>N</i> = 223
Age (years)			
Mean (SD)	44.9 (12.58)	46.4 (11.14)	45.7 (11.89)
Time since diagnosis (years)			
Mean (SD)	12.9 (10.35)	11.1 (10.02)	12.0 (10.21)
Sex, <i>n</i> (%)			
Male	39 (34.2)	46 (42.2)	85 (38.1)
Female	75 (65.8)	63 (57.8)	138 (61.9)
Race, <i>n</i> (%)			
White	106 (93.0)	102 (93.6)	208 (93.3)
Black or African American	6 (5.3)	5 (4.6)	11 (4.9)
Asian	1 (0.9)	1 (0.9)	2 (0.9)
Multiple	1 (0.9)	1 (0.9)	2 (0.9)
Ethnicity, <i>n</i> (%)			
Not Hispanic or Latino	108 (94.7)	99 (90.8)	207 (92.8)
Hispanic or Latino	5 (4.4)	7 (6.4)	12 (5.4)
Not reported	0 (0.0)	1 (0.9)	1 (0.4)
Unknown	1 (0.9)	2 (1.8)	3 (1.3)

AD antidepressant, SD standard deviation

**Table 2** PHQ-9 and MADRS total score and change from baseline score at days 15 and 28

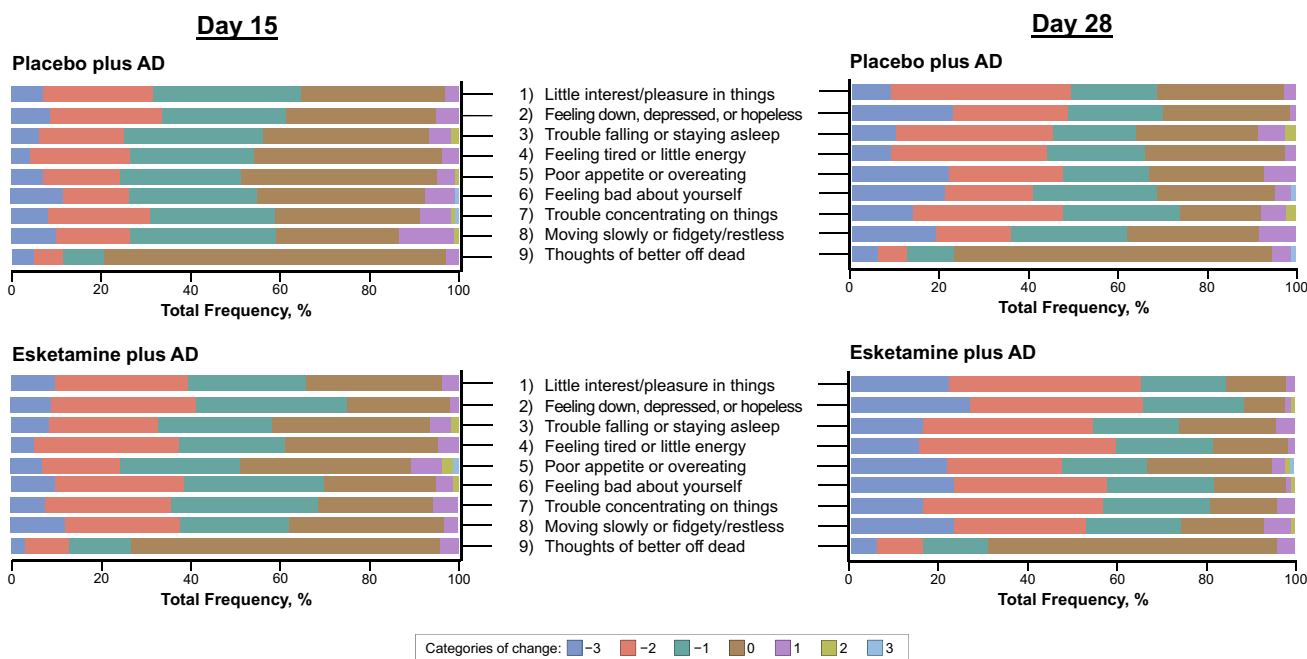
Time point	Absolute score				Change from baseline		Treatment difference		
	Placebo plus AD		Esketamine plus AD		Placebo plus AD	Esketamine plus AD	Mean (SE)	95% CI	p value
	n	Mean (SD)	N	Mean (SD)	Mean (SD)	Mean (SD)			
<b>PHQ-9</b>									
Baseline	109	20.4 (3.73)	114	20.2 (3.63)					
Day 15	104	13.2 (7.20)	111	11.2 (6.33)	- 7.1 (6.87)	- 9.0 (6.45)	- 1.8 (0.91)	- 3.62 to - 0.04	0.045
Day 28	100	10.2 (7.68)	104	7.3 (5.74)	- 10.2 (7.80)	- 13.0 (6.42)	- 2.8 (1.00)	- 4.75 to - 0.81	0.006
<b>MADRS</b>									
Baseline	109	37.3 (5.66)	114	37.0 (5.69)					
Day 15	102	27.2 (11.37)	107	24.8 (10.06)	- 10.0 (11.63)	- 12.1 (10.58)	- 2.0 (1.54)	- 5.06 to 1.00	0.189
Day 28	100	20.6 (12.70)	101	15.5 (10.67)	- 17.0 (13.88)	- 21.4 (12.32)	- 4.4 (1.85)	- 8.10 to - 0.80	0.017

AD antidepressant, MADRS Montgomery-Asberg Depression Rating Scale, PHQ-9 Patient Health Questionnaire 9-item, SE standard error

### 3.1 PHQ-9 Patient-Reported Outcome Findings

Total scores on the PHQ-9 improved from baseline to each post-baseline time point in both treatment groups (Table 2). However, the magnitude of change in PHQ-9 was larger in the esketamine plus AD arm compared with the placebo plus AD arm at both day 15 (- 1.8-point mean difference [SE 0.91] between arms; *p* = 0.045 [95% CI - 3.62 to - 0.04]) and day 28 (- 2.8-point mean difference [SE 1.00] between arms; *p* = 0.006 [95% CI - 4.75 to - 0.81]).

Similarly, distributions of score change at the item level demonstrate that most patients experienced improvement (i.e., a negative change score) on all items except Item 9 (“thoughts you would be better off dead”) from baseline to both day 15 and day 28 (Fig. 1) in both treatment groups. Notably, the large proportion of patients in both groups reporting no change on this item may be a consequence of few patients indicating high levels of suicide risk at baseline. Indeed, the study excluded patients considered to be at serious risk for suicide, as determined through other



**Fig. 1** Proportions of categorical change from baseline to day 15 and day 28 of Patient Health Questionnaire 9-item (PHQ-9) items by treatment group. The PHQ-9 is a patient-reported outcome meas-

ure used to assess depressive symptoms, with each item rated on a 4-point scale. Higher scores indicate greater severity; negative change scores represent improvement. AD antidepressant

methods. However, the improvement was greater in the esketamine plus AD group compared with the placebo plus AD group across all items, especially at day 28.

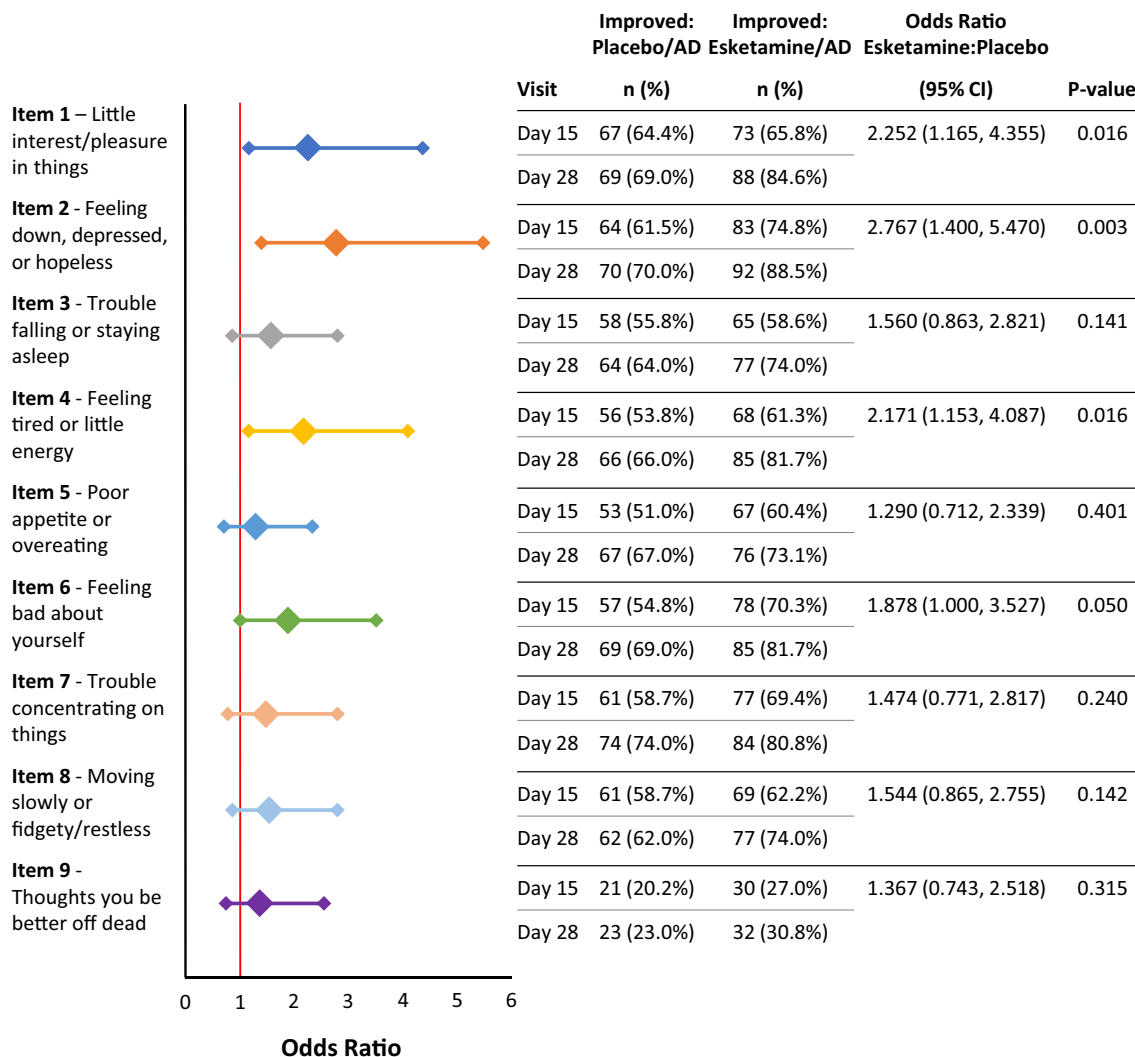
The proportion of patients who experienced a  $\geq 1$ -point improvement was greater in the esketamine plus AD group compared with placebo plus AD for all items on the PHQ-9 at both day 15 and day 28 (Fig. 2). Furthermore, the likelihood of experiencing any improvement over the course of the study was numerically larger in the esketamine plus AD arm compared with placebo plus AD for all nine PHQ-9 items, but particularly for four of the nine PHQ-9 items (for which the nominal  $p \leq 0.05$ ):

- Item 1—“Little interest/pleasure in things”: OR 2.252 (95% CI 1.165–4.355)

- Item 2—“Feeling down, depressed, or hopeless”: OR 2.767 (95% CI 1.400–5.470)
- Item 4—“Feeling tired or having little energy”: OR 2.171 (95% CI 1.153–4.087)
- Item 6—“Feeling bad about yourself – or that you are a failure or have let yourself or your family down”: OR 1.878 (95% CI 1.000–3.527).

### 3.2 MADRS Clinical-Reported Outcome Findings

Total scores on the MADRS improved from baseline to each post-baseline time point in both treatment groups (Table 2). However, the magnitude of change in MADRS was numerically larger in the esketamine/oral AD treatment arm compared with placebo plus AD at day 15 (– 2.0-point mean difference [SE 1.54] between arms;  $p = 0.189$  [95%



**Fig. 2** Distributions forest plot of odds ratios for improvement in Patient Health Questionnaire 9-item (PHQ-9). The odds ratio represents the likelihood of improving over the course of the study and includes both day 15 and day 28 data. AD antidepressant, CI confidence interval

CI – 5.06 to 1.00]) and statistically significantly larger at day 28 (– 4.4-point mean difference [SE 1.85] between arms;  $p = 0.017$  [95% CI – 8.10 to – 0.80]).

Similarly, distributions of score change at the item level demonstrate that most patients experienced improvement (i.e., a negative change score) or remained stable on all items from baseline to Day 15 in both treatment groups (Fig. 3). The day 28 distribution of score change at the item level demonstrates that most patients experience improvement on all items in both treatment groups (Fig. 3). However, the proportions of those who improved were greater in the esketamine plus AD group compared with the placebo plus AD group across all items, especially at day 28.

More patients in the esketamine plus AD group were rated as having improved by at least 2 points compared with placebo plus AD for all items on the MADRS at both day 15 and day 28 (Fig. 4). Furthermore, the likelihood of experiencing any improvement over the course of the study was larger in the esketamine plus AD treatment arm compared with the placebo plus AD arm for five of the ten MADRS items (nominal  $p < 0.05$ ):

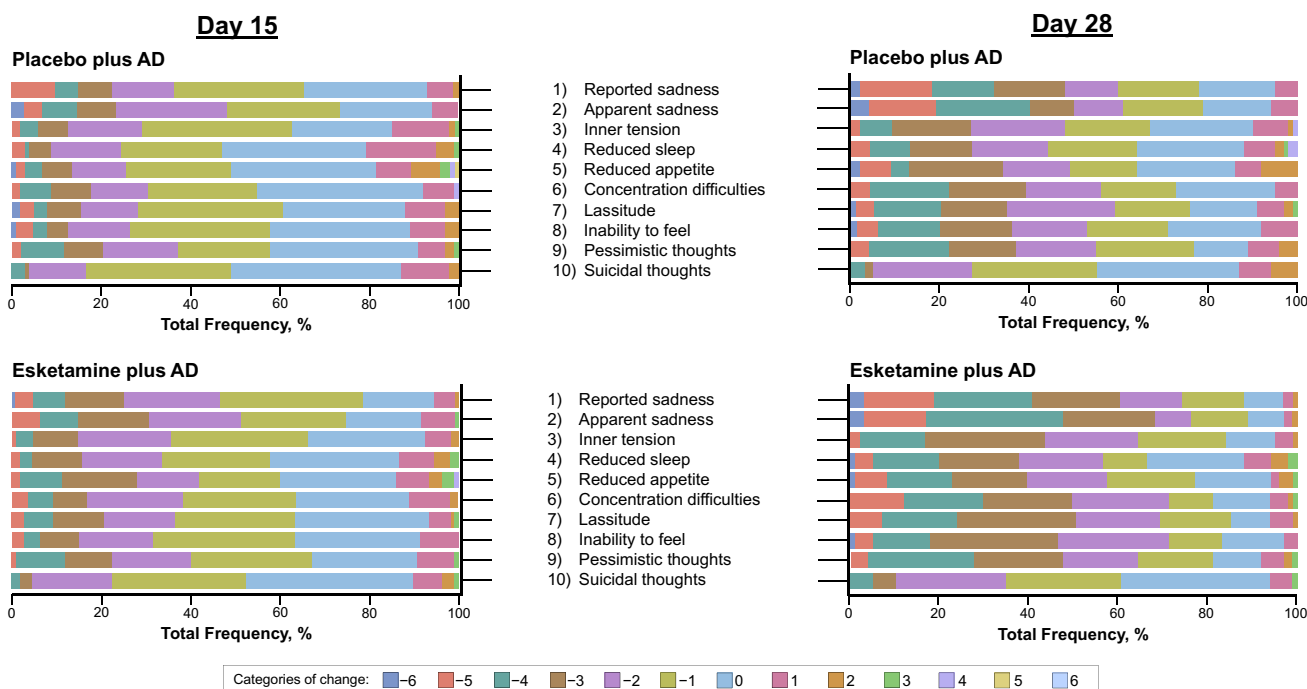
- Item 1—“Reported sadness”: OR 1.844 (95% CI 1.014–3.354)
- Item 2—“Apparent sadness”: OR 2.007 (95% CI 1.096–3.674)

- Item 3—“Inner tension”: OR 1.891 (95% CI 1.080–3.313)
- Item 6—“Concentration difficulties”: OR 1.880 (95% CI 1.054–3.354)
- Item 8—“Inability to feel”: OR 2.099 (95% CI 1.180–3.735).

### 4 Discussion

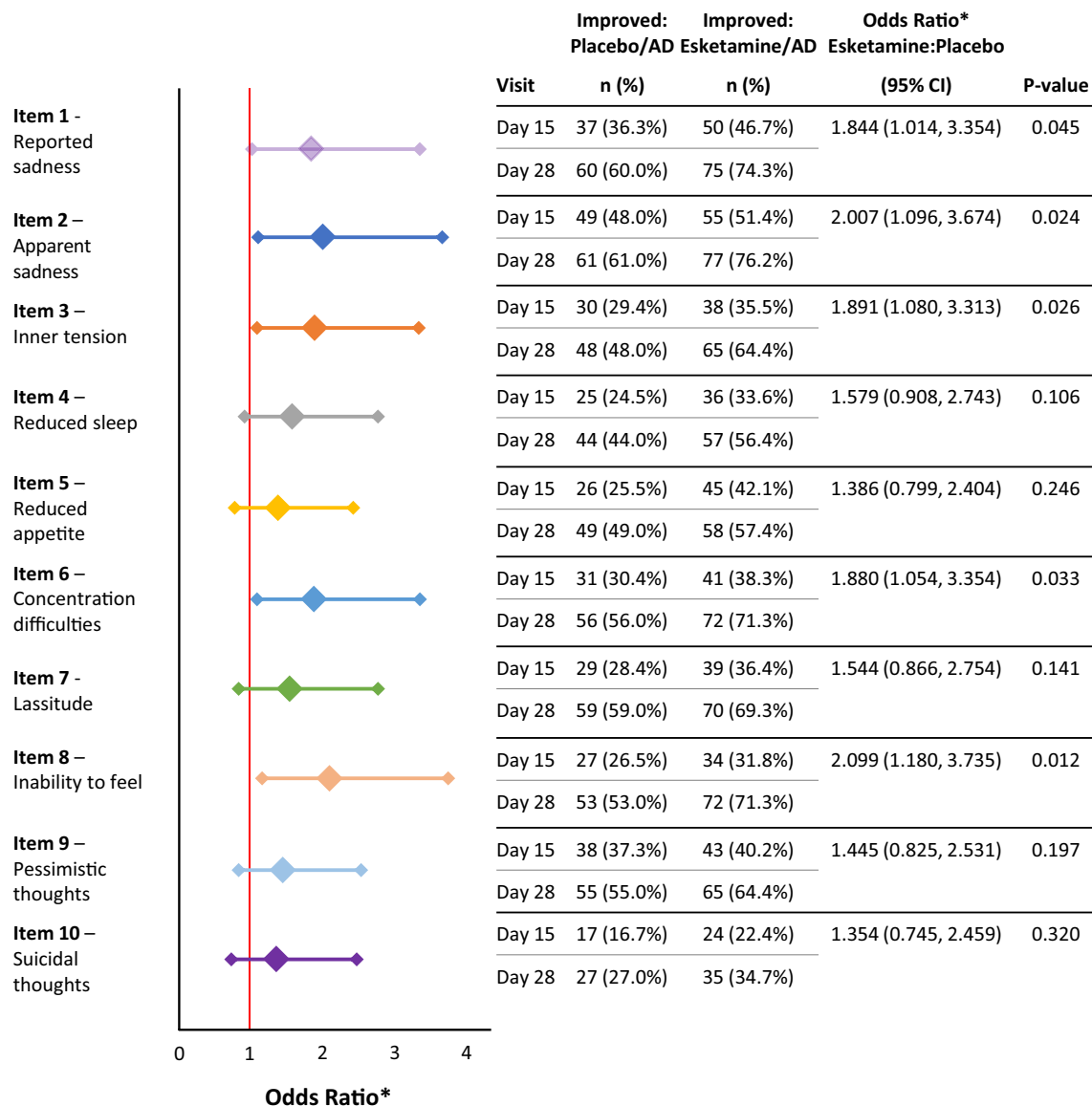
This study demonstrates a pattern of item-level results congruent with the total PHQ-9 and MADRS score change, favoring treatment with esketamine nasal spray plus oral AD versus oral AD plus intranasal placebo. Four out of the nine items on the PHQ-9 and five of the ten items on the MADRS achieved statistical significance, providing a detailed account of the magnitude of which specific depressive symptoms, as measured by single items, are likely to improve from treatment with esketamine nasal spray.

In this study, there was a favorable response to all items on both the PHQ-9 and MADRS, indicating overall improvement in depression for those treated with esketamine nasal spray plus oral AD. While the change in total score is an important indicator of change in depression symptom severity, understanding the specific items contributing more, or less, to the total score change is helpful in interpreting



**Fig. 3** Proportions of categorical change from baseline to day 15 and day 28 of Montgomery-Asberg Depression Rating Scale (MADRS) items by treatment group. The MADRS is a 10-item clinician-rated

scale used to measure depression severity with each item rated on a 7-point scale. Higher scores indicate a more severe condition; negative change scores represent improvement. AD antidepressant



**Fig. 4** Montgomery-Asberg Depression Rating Scale (MADRS) items from baseline to days 15/28 in esketamine plus AD versus placebo plus AD. \*The odds ratio represents the likelihood of improv-

ing over the course of the study and includes both day 15 and day 28 data. An odds ratio > 1 favors esketamine over placebo. AD antidepressant, CI confidence interval

treatment efficacy. Each item represents symptoms of depression which are important diagnostic criteria. By isolating the results of each item, it is possible to determine which symptoms are more likely to improve, enhancing patient understanding of potential treatment effects. For example, based on these results, a clinician can explain that the odds of not feeling down, depressed, or hopeless may be twice that for those treated with esketamine plus AD compared with AD plus placebo after 1 month of treatment. As part of the shared decision-making process, patients and clinicians can weigh the likelihood of individual symptom improvement against treatment risks [16].

The particularly favorable results in PHQ-9 items 1, 2, 4, and 6 and MADRS items 1, 2, 3, 6 and 8 suggest improved efficacy of esketamine/oral AD over placebo plus AD in these specific symptoms among TRD adult patients. PHQ-9 items 1 and 2 (“Little interest/pleasure in doing things” and “Feeling down, depressed, or hopeless”) are both considered cardinal symptoms of depression; a patient will be diagnosed as depressed only if endorsing at least one of these two items. As all items are important for diagnostic criteria, patients and clinicians can see what other items are driving overall score response for those on esketamine plus AD compared with those on placebo plus AD by highlighting items with statistically significant



improvement. On the PHQ-9, these were items 4 and 6 (“Feeling tired or having little energy” and “Feeling bad about yourself – or that you are a failure or have let yourself or your family down”). On the MADRS, reported sadness, apparent sadness, inner tension, concentration difficulties and the inability to feel (anhedonia), statistically significantly improved. Together, these results inform the ways in which patients experience clinical improvement with higher specificity than is conveyed by the total score [17].

Of note, the likelihood of improving was the greatest (OR > 2.0), relative to the other items, for “Little interest/pleasure in things”, “Feeling down, depressed, or hopeless”, and “Feeling tired or little energy” on the PHQ-9. Similarly, on the MADRS, the likelihood of improving was greatest (OR > 2.0) on the items measuring apparent sadness (item 2) and anhedonia (item 8). On both instruments, the symptom that was least likely to improve (other than suicidal thoughts) was having a reduced appetite, as measured by item 5 for both the PHQ-9 and MADRS.

One limitation of the study was the short follow-up time; measurements were taken at baseline and two other time points during the double-blind treatment period. The 4-week duration of the induction phase was chosen to provide sufficient time for the onset of efficacy in the oral AD + placebo group [18]. Meta-analysis results suggest that treatment difference is consistent for trials of 4–8 weeks’ duration, suggesting a duration of 4 weeks is sufficient [19, 20]. Another limitation is the absence of evaluation of the oral AD used. Clinicians prescribed one of four oral ADs for the concomitant treatment; however, any differences of each AD were not analyzed. Moreover, since these are results from one randomized trial, further studies are needed to replicate these findings.

## 5 Conclusions

Single item analysis of individual items allows for a detailed, comprehensive understanding of impact of treatment on patient symptoms as well as magnitude of response. Treatment with esketamine plus AD can lead to improvement in TRD patients based on both clinician and patient evaluation of symptoms. This analysis shows the likelihood of improving the individual items on the PHQ-9 and MADRS, enhancing the interpretability of the overall score change for patients and clinicians. For the PHQ-9, the greatest improvements were on items representing cardinal symptoms of depression; examination of the individual questions on the PHQ-9 highlight the clinical and patient relevance of these results. For the MADRS, the greatest improvement was seen on inability to feel, as reported by clinicians [21].

**Acknowledgements** Copy editing and proofreading services were provided by Chad Green and David Navarro of Clinical Outcomes Solutions, and Heather Rozjabeck of Janssen Research and Development LLC. Funding was provided by Janssen Research and Development LLC.

## Declarations

**Funding** This study and its publication was sponsored by Janssen Pharmaceutical Companies of Johnson & Johnson.

**Potential conflicts of interests** CJ, VP, WD, KC, and JS are employees of Janssen Pharmaceutical Companies of Johnson & Johnson. LF and SH declare no competing interests.

**Availability of data and materials** The datasets generated and/or analyzed during the current study are not publicly available, but are available from the corresponding author on reasonable request.

**Author contributions** All authors contributed to the conception and design, analysis, interpretation, manuscript revisions, approval of the final version to be published, and agree to be accountable for all aspects of the work.

**Ethics approval** All study procedures were in accordance with ethical standards of the 1964 Helsinki Declaration and its later amendments, relevant laws, institutional guidelines, and approved by an institutional review board.

**Consent to participate** All participants provided informed consent.

**Consent for publication** Not applicable.

**Code availability** Not applicable.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

## References

1. Vos T, Abajobir AA, Abate KH, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: a systematic analysis for the global burden of disease study 2016. *Lancet*. 2017;390(10100):1211–59.
2. Al-Harbi KS. Treatment-resistant depression: therapeutic trends, challenges, and future directions. *Patient Prefer Adher*. 2012;6:369.
3. Fekadu A, Wooderson SC, Markopoulo K, Donaldson C, Papadopoulos A, Cleare AJ. What happens to patients with

- treatment-resistant depression? A systematic review of medium to long term outcome studies. *J Affect Disord.* 2009;116(1–2):4–11.
4. Johnston KM, Powell LC, Anderson IM, Szabo S, Cline S. The burden of treatment-resistant depression: a systematic review of the economic and quality of life literature. *J Affect Disord.* 2019;242:195–210.
  5. Edwards S, Hamilton V, Nherera L, Trevor N. Lithium or an atypical antipsychotic drug in the management of treatment-resistant depression: a systematic review and economic evaluation. *Health Technol Assess.* 2013;17(54):1.
  6. Daly E, Singh J, Fedgchin M, et al. Efficacy and safety of intranasal esketamine adjunctive to oral antidepressant therapy in treatment-resistant depression: a randomized clinical trial. *JAMA Psychiat.* 2018;75(2):139–48.
  7. Fedgchin M, Trivedi M, Daly E, et al. Efficacy and safety of fixed-dose esketamine nasal spray combined with a new oral antidepressant in treatment-resistant depression: results of a randomized, double-blind, active-controlled study (TRANSFORM-1). *Int J Neuropsychopharmacol.* 2019;22(10):616–30.
  8. Popova V, Daly E, Trivedi M, et al. Efficacy and safety of flexibly dosed esketamine nasal spray combined with a newly initiated oral antidepressant in treatment-resistant depression: a randomized double-blind active-controlled study. *Am J Psychiatry.* 2019;176(6):428–38.
  9. Ochs-Ross R, Daly E, Zhang Y, et al. Efficacy and safety of esketamine nasal spray plus an oral antidepressant in elderly patients with treatment-resistant depression TRANSFORM-3. *Am J Geriatr Psychiatry.* 2020;28(2):121–41.
  10. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med.* 2001;16(9):606–13.
  11. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th ed. Arlington: American Psychiatric Association; 2013.
  12. Hudgens S, Floden L, Symonds T, Slagle A. PRM9—a longitudinal, item-level analysis methodology to support the interpretability of multi-item patient-reported outcomes. *Value Health.* 2018;21:S357.
  13. Repo J, Rosqvist E. Guidelines for translation and cross-cultural adaptation of non-technical skills rating instruments. *Value Health.* 2016;19(7):A358.
  14. Hudgens S, Floden L, Blackowicz M, et al. Meaningful change in depression symptoms assessed with the Patient Health Questionnaire (PHQ-9) and Montgomery-Åsberg Depression Rating Scale (MADRS) among patients with treatment resistant depression in two, randomized, double-blind, active-controlled trials of esketamine nasal spray combined with a new oral antidepressant. *J Affect Disord.* 2021;281:767–75.
  15. Montgomery SA, Möller H-J. Is the significant superiority of escitalopram compared with other antidepressants clinically relevant? *Int Clin Psychopharmacol.* 2009;24(3):111–8.
  16. Stiggelbout AM, Van der Weijden T, De Wit MP, Frosch D, Légaré F, Montori VM, Trevena L, Elwyn G. Shared decision making: really putting patients at the centre of healthcare. *BMJ.* 2012;344.
  17. El-Khalili N, Joyce M, Atkinson S, et al. Extended-release quetiapine fumarate (quetiapine XR) as adjunctive therapy in major depressive disorder (MDD) in patients with an inadequate response to ongoing antidepressant treatment: a multicentre, randomized, double-blind, placebo-controlled study. *Int J Neuropsychopharmacol.* 2010;13(7):917–32.
  18. European Medicines Agency. Guideline for the clinical investigation of medicinal products in the treatment of depression (EMA/CHMP/185423/2010 REV.2) (2013)
  19. Khin NA (2013) Update on FDA exploratory analyses of aggregated efficacy data from antidepressant trials. Paper presented at: NCDEU regulatory plenary. New FDA & EMA initiatives in Alzheimer's disease, depression and schizophrenia. May 30, 2013, Phoenix AZ
  20. Yang P (2013) Recent advances in design and methodology in psychiatric clinical trials. Paper presented at: joint statistical meetings; August 3–8, 2013; Montréal, QC, Canada
  21. Ballard E, Wills K, Lally N, et al. Anhedonia as a clinical correlate of suicidal thoughts in clinical ketamine trials. *J Affect Disord.* 2017;218:195–200.

## Authors and Affiliations

Lysbeth Floden<sup>1</sup>  · Stacie Hudgens<sup>1</sup> · Carol Jamieson<sup>2</sup> · Vanina Popova<sup>3</sup> · Wayne C. Drevets<sup>4</sup> · Kimberly Cooper<sup>5</sup> · Jaskaran Singh<sup>4,6</sup>

✉ Lysbeth Floden  
libby.floden@clinoutsolutions.com

<sup>1</sup> Clinical Outcomes Solutions, 1820 E River RD, Suite 220, Tucson, AZ 85718, USA

<sup>2</sup> Janssen Research and Development LLC, Milpitas, CA, USA

<sup>3</sup> Janssen Research and Development LLC, Beerse, Belgium

<sup>4</sup> Janssen Research and Development LLC, San Diego, CA, USA

<sup>5</sup> Janssen Research and Development LLC, Spring House, PA, USA

<sup>6</sup> Present Address: Neurocrine Biosciences, San Diego, CA, USA