Demonstration of the Relationships Among Clinical Global Impression of Severity of Depression Scale and Montgomery-Åsberg Depression Rating, Patient Health Questionnaire-9, and Sheehan Disability Scales

Turkoz I,[1] Alphs L,[2] Singh J,[1] Lim P,[1] Lane R,[1] DiBernardo A,[2] Shawi M,[2] Hough D[1]

[1]Janssen Research and Development, LLC, Titusville, NJ, USA; [2]Janssen Scientific Affairs, LLC, Titusville, NJ, USA

INTRODUCTION

- The Montgomery-Åsberg Depression Rating Scale (MADRS)[1] is widely used to assess mood symptoms in clinical trials of patients with major depression and treatment-resistant depression (TRD)
  - Data derived from the MADRS are frequently expressed as changes in total scale scores or as changes relative to predefined threshold scores
  - A limitation of this approach is that there is no well-established definition of a clinically meaningful change or clinically significant change for this scale when viewed at the individual-patient level
  - The usual MADRS cutoff points are[2-4]:
    - 0 to 6, normal/symptoms absent
    - 7 to 19, mild depression
    - 20 to 34, moderate depression
    - >34, severe depression
- The Sheehan Disability Scale (SDS)[5] and the Patient Health Questionnaire (PHQ-9) for depression[6] are brief patient-reported tools used to assess functional impairment and to assess depression, respectively, in patients with major depression
  - As with the MADRS, clinically meaningful changes in functioning as measured by the SDS or the PHQ-9 are not well established
  - Clinicians and researchers have different values for clinically meaningful changes and clinically significant changes
  - Functional remission has been defined as SDS ≤6 at end point,[7] and a PDQ-9 score of <5 is considered to be "remission"[8]
  - Potential PHQ-9 cutoffs are (available at http://www.phqscreeners.com/)
    - 0-4, none or minimal
    - 5-9, mild
    - 10-14, moderate
    - 15-19, moderately severe
    - 20-27, severe
- The Clinical Global Impressions–Severity (CGI-S) scale[9] for depression is widely used to measure clinically meaningful symptomatic changes. A change of at least 1 point on the CGI-S can be used as a proxy for clinical meaningfulness, and a change of at least 2 points on the CGI-S can be used as a proxy for clinical significance
- The present analysis explores the relationship among ratings on the MADRS, SDS, and PHQ-9 and the CGI-S measured in the same subjects by determining change scores on the MADRS, SDS, and PHQ-9 that correspond to a 1- and 2-point change on the CGI-S

METHODS

- This post hoc analysis used an international clinical trial database (N=223) composed of data from a 4-week, randomized, active-controlled study (NCT02418585) of antidepressant therapy versus active comparator in symptomatic subjects with major depressive disorder who had not benefited from ≥2 prior pharmacologic treatments for their TRD within the current episode
- All subjects were rated with the CGI-S (at baseline and on days 4, 8, 11, 15, 22, and 28); MADRS (at baseline and on days 2, 8, 15, 22, and 28); PHQ-9 (at baseline and on days 15 and 28); and SDS (at baseline and on days 15 and 28)
- Parametric and nonparametric simple and multiple regression models (with explanatory study design variables for treatment, country, class of antidepressant [serotonin and norepinephrine reuptake inhibitors or selective serotonin reuptake inhibitors], and baseline score) were used to explore relationships between MADRS, PHQ-9, and SDS ratings with CGI-S score from baseline to the week-4 end point
- Proxy for clinically meaningful improvement was defined as a 1-point change in the CGI-S. A 2-point change on the CGI-S was defined as a clinically significant change
- Assumptions such as a linear relationship between variables and equal variances of error terms were evaluated. Regression coefficients and r^2 were assessed at each visit to illustrate consistency of observed estimates
- Ordinal logistic regression models were used to determine expected values of MADRS, PHQ-9, and SDS at end point for each score of the CGI-S
- Additional logistic regression models were fit using the change score of the CGI-S as a categorical variable. The categories of change in the CGI-S were defined as improvement (≤−3, −2, or −1); unchanged (0); or worsening (≥1)
- Linking analyses were performed to determine the correspondence of the CGI-S score to the MADRS, PHQ-9, and SDS scores, by mapping percentiles of the cumulative distribution of observed values for the 2 scales at baseline and each subsequent measurement by the method of Leucht[10]
- No adjustment was made for multiplicity

Supported by Janssen Scientific Affairs, LLC

RESULTS

- Baseline demographics are shown in **Table 1**

Table 1. Baseline demographics

Parameter	Overall (N = 223)
Age, y, mean (SD)	45.7 (11.9)
Sex, n (%)	
Male	85 (38.1)
Female	138 (61.9)
Race, n (%)	
Caucasian	208 (93.3)
Black or African American	11 (4.9)
Asian	2 (0.9)
Other	2 (0.9)
Region, n (%)	
Europe	134 (60.1)
North America	89 (39.9)
Class of antidepressants, n (%)	
SNRI	152 (68.2)
SSRI	71 (31.8)
Duration of current episode, weeks, mean (SD)	114.6 (158.0)
Age when diagnosed with major depressive disorder, y, mean (SD)	33.7 (12.9)
CGI-S, mean (SD)	5.1 (0.8)
MADRS total score, mean (SD)	37.1 (5.7)
PHQ-9 total score, mean (SD)	20.3 (3.7)
SDS, mean (SD)	24.1 (4.2)

SNRI, serotonin and norepinephrine reuptake inhibitor; SSRI, selective serotonin reuptake inhibitor.

- Results from simple and multiple regression model results are shown in **Table 2**
  - With multiple regression, a 1-point change in the CGI-S corresponded to a 6.4-point (SE=0.4) change from baseline in the MADRS; a 3.1-point (SE=0.2) change from baseline in the PHQ-9; and a 3.8-point (SE=0.3) change from baseline in the SDS

Table 2. MADRS total, PHQ-9 total, and SDS total change from baseline to end point during double-blind induction phase associated with a 1-point change in CGI-S

	MADRS total change	PHQ-9 total change	SDS total change
Simple regression			
Estimate (SE)	7.0 (0.4)	3.1 (0.3)	3.9 (0.3)
P value	<0.001	<0.001	<0.001
r^2	0.543	0.384	0.459
Multiple regression			
Estimate (SE)	6.4 (0.4)	3.1 (0.2)	3.8 (0.3)
P value	<0.001	<0.001	<0.001
r^2	0.637	0.544	0.521

- Few subjects had a CGI-S change score ≤−3 or >1. Although the linearity of the relationship between dependent and independent variables was examined using various graphic techniques, having insufficient subjects for these score ranges may have severely biased the linearity assumption
- Proportion of subjects meeting proxy clinically meaningful change at the earliest measurable time and proportion of subjects meeting proxy clinically significant change at the latest measurable time are shown in **Table 3**

Table 3. Proportion of subjects demonstrating clinically meaningful change and clinically significant change in MADRS, PHQ-9, and SDS scales at earliest measurable time and at end point

Scale		Oral AD + PBO n (%)	Oral AD + Study Drug n (%)
MADRS			
6-point change at day 2	No	61 (59.8)	42 (38.5)
	Yes	41 (40.2)	67 (61.5)
12-point change at end point	No	45 (41.3)	35 (31.3)
	Yes	64 (58.7)	77 (68.8)
PHQ-9			
3-point change at day 15	No	32 (30.8)	23 (20.7)
	Yes	72 (69.2)	88 (79.3)
6-point change at end point	No	33 (31.4)	20 (18.0)
	Yes	72 (68.6)	91 (82.0)
SDS			
4-point change at day 15	No	21 (47.7)	15 (32.6)
	Yes	23 (52.3)	31 (67.4)
8-point change at end point	No	42 (47.7)	25 (26.9)
	Yes	46 (52.3)	68 (73.1)

AD, antidepressant; PBO, intranasal placebo.

- A proportional odds model was used to examine the relationship between CGI-S scores at end point and MADRS scores (**Table 4**), PHQ-9 scores (**Table 5**), and SDS scores (**Table 6**) at end point

Table 4. Distribution of MADRS scores by CGI-S score at end point (probabilities ≥20% are highlighted)

	Predictive probability of MADRS total score at end point									
	0-6 (n = 42)		7-12 (n = 45)		13-19 (n = 37)		20-34 (n = 63)		>34 (n = 34)	
CGI-S	n	%	n	%	n	%	n	%	n	%
1 = normal	13	31.0	4	8.9	0	0	0	0	0	0
2 = minimally ill	15	35.7	18	40.0	9	24.3	0	0	0	0
3 = mildly ill	12	28.6	18	40.0	22	59.5	8	12.69	0	0
4 = moderately ill	2	4.8	4	8.9	6	16.21	39	61.9	8	23.5
5 = markedly ill	0	0	1	2.2	0	0	15	23.8	19	55.9
6, 7 = severely ill, extremely ill	0	0	0	0	0	0	1	1.58	7	20.6

Ordinal logistic regression model with MADRS as a predictor (OR = 1.3; 95% CI, 1.2 to 1.3). On average, there was 25% probability of a shift in one CGI-S category at end point when the MADRS score changed by 1 unit.

Table 5. Distribution of PHQ-9 scores by CGI-S score at end point (probabilities ≥20% are highlighted)

	Predictive probability of PHQ-9 total score at end point									
	0-4 (n = 67)		5-9 (n = 72)		10-14 (n = 26)		15-19 (n = 23)		20-27 (n = 28)	
CGI-S	n	%	n	%	n	%	n	%	n	%
1 = normal	10	14.9	7	9.7	0	0	0	0	0	0
2 = minimally ill	28	41.8	13	18.1	1	3.8	0	0	0	0
3 = mildly ill	23	34.3	32	44.4	5	19.2	0	0	0	0
4 = moderately ill	6	9.0	16	22.2	17	65.4	11	47.8	7	25.0
5 = markedly ill	0	0	4	5.6	3	11.5	12	52.2	16	57.1
6, 7 = severely ill, extremely ill	0	0	0	0	0	0	0	0	5	17.9

Ordinal logistic regression model with PHQ-9 as a predictor (OR = 1.4; 95% CI, 1.3 to 1.5). On average, there was 39% probability of a shift in one CGI-S category at end point when the PHQ-9 score changed by 1 unit.

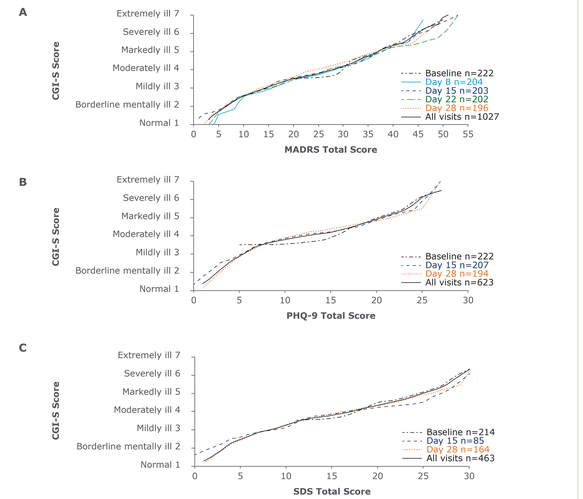
Table 6. Distribution of SDS scores by CGI-S score at end point (probabilities ≥20% are highlighted)

	Predictive probability of SDS total score at end point							
	0-4 (n = 39)		5-8 (n = 28)		9-14 (n = 40)		>14 (n = 75)	
CGI-S	n	%	n	%	n	%	n	%
1 = normal	11	28.2	1	3.6	1	2.5	0	0
2 = minimally ill	13	33.3	9	32.1	5	12.5	3	4.0
3 = mildly ill	14	35.9	11	39.3	16	40.0	9	12.0
4 = moderately ill	1	2.6	6	21.4	13	32.5	30	40.0
5 = markedly ill	0	0	1	3.6	5	12.5	28	37.3
6, 7 = severely ill, extremely ill	0	0	0	0	0	0	5	6.7

Ordinal logistic regression model with SDS as a predictor (OR = 1.3; 95% CI, 1.2 to 1.3). On average, there was 27% probability of a shift in one CGI-S category at end point when the SDS score changed by 1 unit.

- Equipercentile linking results between the CGI-S and the MADRS, PHQ-9, and SDS total scores at each time point and for the combined data set are shown in **Figure 1A-C**
  - The relationship between the CGI-S score and the MADRS, PHQ-9, and SDS total scores followed a linear trend, and the relationship between the 2 variables at each visit showed consistency with increasing correlation coefficients over time

Figure 1. Relationship between CGI-S and MADRS total scores (A), PHQ-9 total scores (B), and SDS total scores (C) by visit using local equipercentile method



DISCUSSION AND CONCLUSIONS

- **Clinically meaningful improvement at a patient level in depressive symptoms identified by the CGI-S scale corresponds to approximately 6.4-, 3.8-, and 3.1-point improvement from baseline on the MADRS, SDS, and PHQ-9, respectively, in adult subjects (18 to 65 years) with TRD**
- **Additional studies are needed to confirm these findings in subjects who demonstrate broader ranges of change and to show that these relationships hold for both clinical worsening and improvement**

REFERENCES

1. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry*. 1979;134:382-389.
2. Herrmann N, Black SE, Lawrence J, Szekely C, Szalai JP. The Sunnybrook Stroke Study: a prospective study of depressive symptoms and functional outcome. *Stroke*. 1998;29(3):618-624.
3. Muller-Thomsen T, Arlt S, Mann U, Mass R, Ganzer S. Detecting depression in Alzheimer's disease: evaluation of four different scales. *Arch Clin Neuropsychol*. 2005;20(2):271-276.
4. McDowell I. *Measuring Health: A Guide to Rating Scales and Questionnaires*. 3rd ed. New York, NY: Oxford University Press, Inc.; 2006.
5. Sheehan DV. *The Anxiety Disease. A Leading Psychiatrist Offers New Hope for Victims of Severe Anxiety*. New York, NY: Charles Scribner & Sons; 1983.
6. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. *JAMA*. 1999;282(18):1737-1744.
7. Sheehan KH, Sheehan DV. Assessing treatment effects in clinical trials with the discan metric of the Sheehan Disability Scale. *Int Clin Psychopharmacol*. 2008;23(2):70-83.
8. Kroenke K, Spitzer RI. The PHQ-9: a new depression diagnostic and severity measure. *Psychiatr Ann*. 2002;32:7.
9. Guy W. ECDEU Assessment Manual for Psychopharmacology (028 Clinical Global Impressions [CGI]). Rockville, MD: NIMH; 1976.
10. Leucht S, Kane JM, Kissling W, Hamann J, Etschel E, Engel RR. What does the PANSS mean? *Schizophr Res*. 2005;79(2-3):231-238.

ACKNOWLEDGMENTS

The authors thank Matthew Grzywacz, PhD, and Lynn Brown, PhD, of ApotheCom (Yardley, PA) for their writing and editorial assistance.

DISCLOSURES

One or more authors report potential conflicts, which are described in the program.