Evidence to Support Montgomery-Asberg Depression Rating Scale Administration Every 24 Hours to Assess Rapid Onset of Treatment Response

Kristen M. Johnson, PhD, MPH[a]; Jacob M. Devine, MA[b]; Kai Fai Ho, PhD[c]; Kellee A. Howard, MA, MSc[b]; Todd L. Saretsky, MPH[b]; and Carol A. Jamieson, BSc, FIBMS[a,*]

ABSTRACT

Objective: This study investigated the suitability of the Montgomery-Asberg Depression Rating Scale (MADRS), with a 24-hour recall period (MADRS-24hr), to assess the rapid onset of the antidepressant effect of a treatment in patients with treatment-resistant depression (TRD). Psychometric properties of the MADRS-24hr were assessed together with qualitative assessment of content validity.

Methods: Content validity was assessed using semistructured interviews conducted from November 2013 to December 2013 in patients (18–64 years old) with TRD who met *DSM-IV* diagnostic criteria and health care professionals (HCPs) experienced in treating major depressive disorder and familiar with using the MADRS. The psychometric properties of MADRS-24hr were evaluated using data from 2 randomized clinical studies involving patients with TRD.

Results: A total of 23 patients (15 [65%] women) with TRD (mean age = 45 years) and 11 HCPs were interviewed. With the exception of reduced sleep, the majority of patients and HCPs reported that the items captured in the MADRS can fluctuate in a 24-hour period. The majority of participants also reported that a meaningful change in depression symptoms could be assessed in a 24-hour recall period, except for reduced sleep and appetite. Assessment of the psychometric properties of the MADRS-24hr showed that this instrument had high internal consistency reliability (Cronbach α of 0.84 and 0.91) and test-retest reliability (intraclass correlation coefficients of 0.96 and 0.91), had construct validity, and was responsive to change following an intervention.

Conclusions: Overall, results suggest that MADRS-24hr can be used to assess the rapid onset of antidepressant efficacy of a treatment in patients with TRD.

Trial Registration: ClinicalTrials.gov identifiers: NCT01627782 and NCT01640080

J Clin Psychiatry 2016;77(12):1681–1686
dx.doi.org/10.4088/JCP.15m10253

[a]Janssen Global Services, LLC, Raritan, New Jersey
[b]Oxford Outcomes, San Francisco, California
[c]STAT-TU Inc, Toronto, Ontario, Canada
**Corresponding author:* Carol A. Jamieson, BSc, FIBMS, Janssen Global Services, LLC, Patient Reported Outcomes, 700 US, 202 Raritan Ave, Raritan, NJ 08869 (cjamies2@its.jnj.com).

Major depressive disorder (MDD) has a lifetime prevalence of 10%–16% in the general adult population.[1,2] Although there is variability in definitions of treatment-resistant depression (TRD), the accepted definition in a regulatory context is "MDD that persists even after at least 2 separate antidepressant treatment regimens at adequate dose and duration in the current episode."[3] Patients with TRD have more comorbidities and greater medical resource utilization[3–5] than MDD patients who respond to treatment. A recent study in the United States found that when compared with MDD, the impact of TRD on resource utilization was substantial, due to a longer duration of depressive episodes and greater rate of therapy utilization.[3]

The clinician-rated Montgomery-Asberg Depression Rating Scale (MADRS)[6] is a well- established instrument and has reasonably good psychometric properties[7,8] for assessing treatment change in clinical trials of antidepressants, typically using a 7-day recall period. There are examples in which administration of the MADRS has been modified to meet research needs, such as using a self-reported questionnaire,[9,10] and versions for interactive voice recognition, telephone administration,[11–13] and exclusion of somatic scales.[14] Additionally, a modified recall version of the instrument has been used in some clinical trials, for example, a randomized clinical trial of quetiapine for bipolar II disorder in which the MADRS was administered with recall ranging between 1 to 3 days, rather than at 7 days.[15] However, given the potential for novel antidepressants with a rapid onset of action, there is a need to either develop new measures to detect this rapid change in symptom severity in antidepressant clinical trial settings or determine if preexisting instruments can be used with a shortened recall period among patients with MDD. To our knowledge, the present study is the first time that the content validity and psychometric properties of the MADRS with a 24-hour recall period (MADRS-24hr) have been evaluated.

This research focused on understanding MDD symptomatology from the perspectives of both patients and health care professionals (HCPs), thereby confirming the content validity of the MADRS-24hr and assessing the psychometric properties of the MADRS-24hr. Specific research objectives included (1) investigating whether or not symptoms of depression, as covered in the 10 MADRS items, can fluctuate in a 24-hour period and (2) assessing the psychometric properties of the MADRS-24hr. Taken together, these would provide evidence supporting the feasibility of using the MADRS-24hr to assess the rapid onset of antidepressant effect at clinical trial assessment points occurring over 24 hours.

METHODS

This evaluation included (1) cognitive interviews in patients with TRD and HCPs who are experienced in treating MDD and are

familiar with using the MADRS and (2) analysis of 2 clinical trial datasets to assess the psychometric properties of the MADRS-24hr.

Clinical Points

- Most major depressive disorder symptoms, as reported by patients and health care professionals, can fluctuate perceptibly within a 24-hour time period.
- Psychometric data for the Montgomery-Asberg Depression Rating Scale with a 24-hour recall period are supportive of its use to assess rapid antidepressant response in a clinical trial setting.

Cognitive Interviews

Using the approach recommended in the Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims,[16] semistructured interviews were conducted to assess the content validity of the MADRS-24hr in patients with TRD, and in HCPs who treat MDD and TRD and are well versed in administering the MADRS. Given that the objective of this research was to evaluate a 24-hour recall period, patients and HCPs were drawn from medical centers that previously participated in an investigational study of ketamine[17] to ensure that the participants had the potential to have experienced rapidly changing symptoms. Ketamine and esketamine are currently being investigated as treatments for TRD with the potential for a rapid onset of response, within hours to days of the first dose.[17] Patients were eligible if they were 18–64 years old; met the *Diagnostic and Statistical Manual of Mental Disorders*, *Fourth Edition*, diagnostic criteria for recurrent MDD without psychotic features; used ≥ 1 antidepressant for at least the prior 3 months and/or had ever been hospitalized for MDD; had inadequate response to 1 antidepressant in the current episode of depression; and had inadequate response to at least 1 other antidepressant either in the current episode or in a previous episode. Patients with a history of alcohol/substance abuse and/or dependence (prior 6 months) or a history/current diagnosis of psychotic/bipolar disorder, mental retardation, or borderline personality disorders were excluded.

Interviews were conducted (November–December 2013) by trained researchers. At the beginning of the interview, patients were administered the Patient Health Questionnaire–9 item (PHQ-9) to assess the severity of depression symptoms[18] to help characterize the study population. Due to the nature of the semistructured interviews, not every participant provided responses to all questions, as some responses were deemed incomplete or incoherent.

The primary objective was to assess if the depressive symptoms measured by the MADRS varied over a period of 24 hours, thereby providing confirmation that it is appropriate to measure changes in these symptoms 24 hours after treatment from both the patients' and HCPs' perspectives. Patients were asked about each of the items in the MADRS to evaluate (1) whether each symptom was constant or varied within a 24-hour period, (2) if assessment of change within 24-hour was feasible and adequate, and (3) the time taken to feel a meaningful improvement from how they felt previously.

Interviews were audiorecorded and data were analyzed using MAXQDA 10 (VERBI GmbH), a professional qualitative analysis software tool that enables a systematic team-based approach to coding and builds validity and reliability into the analysis. MAXQDA uses a combined thematic and content analytic approach that allows identification of broader concepts while simultaneously examining responses to particular questions or topics.[19] Researchers trained in qualitative methodology conducted the analysis with MAXQDA. A codebook was created that included the code, and examples, and instructions on application of the codes were provided.

Psychometric Analysis

To complement the qualitative data collected through interviews, psychometric properties of the MADRS-24hr were analyzed using data from 2 double-blind, randomized studies that enrolled patients with TRD (Study A,[17] NCT01640080; Study B,[20] NCT01627782). Specifically, this analysis looked at item-level characteristics, scale structure, reliability, validity, responsiveness to change, and minimum important change (MIC). The schedule of assessments for these 2 studies together with key inclusion and exclusion criteria can be found in Table 1.

Patient- and Clinician-Reported Assessments

Studies A and B included patient- and clinician-reported assessments that were used for the psychometric analysis. The Clinical Global Impression-Severity (CGI-S) scale[21] provides an overall clinician-determined summary of disease severity. Similarly, the Clinical Global Impression-Improvement (CGI-I) scale is used to assess patients' improvement/deterioration.[21]

The patient-reported measures analogous to the CGI-S and CGI-I are the Patient Global Impression-Severity (PGI-S) and the Patient Global Impression-Change (PGI-C). Two versions of PGI-S were used (Study A: 4-response options; Study B: 10-response options), with a higher score indicating more severity.[22,23]

The Quick Inventory of Depressive Symptomatology–Self-Report, 16 item version (QIDS-SR_{16}) and a shorter 14-item version with a 24-hour recall period (QIDS-SR_{14}) are patient-completed questionnaires used to measure the overall severity of depressive symptoms[24,25]; the score ranges from 0–27, with a higher score indicating greater symptom severity.

Analytic Approach

Descriptive statistics were used to assess measurement properties of the MADRS-24hr, including evaluation of floor and ceiling effects. The unidimensional structure of

Table 1. Abbreviated Time and Events Schedule for Study A[a] and Study B

	Study A[b]				Study B[b]			
					2 Times per Week		3 Times per Week	
	Day 1	Day 2	Day 3	Day 4	Day 1	Day 4	Day 1	Day 3
Study drug administration	✓			✓	✓	✓	✓	✓
MADRS (7-day recall)	✓				✓		✓	
MADRS (24-hour recall)		✓	✓	✓		✓		✓
CGI-S, PGI-S	✓	✓	✓	✓	✓	✓	✓	✓
CGI-I, PGI-C		✓	✓	✓		✓		✓
QIDS-SR_{16} (7-day recall)	✓				✓		✓	
QIDS-SR_{14} (24-hour recall)		✓	✓	✓		✓		✓

[a]Patient-reported outcomes were administered pre-dose or at a similar time at days 2 and 3 in Study A.

[b]Study inclusion criteria (both Study A and Study B) : age 18–64 years, *DSM-IV-TR* diagnosis of major depressive disorder, 1 inadequate treatment response in the current episode, and at least 1 other inadequate treatment response in either the current episode or a previous episode. Exclusion criteria: comorbid generalized anxiety disorder, panic disorder, obsessive-compulsive disorder, posttraumatic stress disorder, anorexia nervosa, or bulimia nervosa; history or current diagnosis of psychotic disorder, bipolar disorder, mental retardation, or borderline personality disorders; mood disorder with postpartum onset; somatoform disorders or chronic fatigue syndrome; and being acutely suicidal or homicidal (ie, in imminent danger with plan to harm oneself, recent suicide attempt per principal investigator's clinical judgment).

Abbreviations: CGI-I = Clinical Global Impression–Improvement, CGI-S = Clinical Global Impression–Severity, MADRS = Montgomery-Asberg Depression Rating Scale, PGI-C = Patient Global Impression-Change, PGI-S = Patient Global Impression-Severity, QIDS-SR = Quick Inventory of Depressive Symptomatology–Self-Report.

Table 2. Qualitative Interviews: Patient Demographic and Clinical Characteristics (N = 23)

Characteristic	Value
Age, y	
Mean (SD)	45 (11.8)
Range	20–60
Women, n (%)	15 (65)
Prior ketamine experience, n (%)	8 (35)
Race, n (%)	
Caucasian/white	21 (91)
African American/black	2 (9)
Employment status, n (%)	
Employed full-time	7 (30)
Employed part-time	6 (26)
Unemployed[a]	10 (43)
Education, n (%)	
Less than high school degree	2 (9)
High school degree/GED	2 (9)
Some college	8 (35)
Associate's degree	1 (4)
Bachelor's degree	7 (30)
Master's degree or higher	3 (13)
PHQ-9 score	
Mean (SD)	12.4 (6.4)
Range	2–26

[a]Includes unemployed, permanently disabled, temporarily disabled, and retired.

Abbreviations: GED = general educational development, PHQ = Patient Health Questionnaire, SD = standard deviation.

the instrument was assessed by conducting a confirmatory factor analysis. Internal consistency reliability was assessed using Cronbach α. These assessments were done at day 4 pre-dose in Study A and at either day 3 or day 4 pre-dose in Study B.

Test-retest reliability (Study A only) was assessed by comparing day 3 and day 4 pre-dose MADRS-24hr scores on 2 different stable populations, defined as those patients having unchanged responses at both assessments on the CGI-S and PGI-S.

Known-groups validity was assessed by examining the mean MADRS-24hr score at day 4 pre-dose (Study A) and day 3 or day 4 pre-dose (Study B) between groups of patients with differing severity, as defined by 2 measures of patient-reported depression severity, the PGI-S and QIDS-SR_{14}. Differences in group means were assessed using an analysis of variance model. Responsiveness was assessed by comparing change scores of patients whose health state did not change to those patients who showed improvement. Two definitions of improvement in health state were evaluated: (1) at least "minimally improved" on CGI-I and (2) at least "improved" on PGI-C. The magnitude of each within-group change was assessed using a paired *t* test, whereas the magnitude of difference in mean change scores between unchanged patients and patients reporting an improvement was assessed using a 2-sample *t* test. Within-group effect sizes (ES) were computed as the ratio of the mean change score to the standard deviation of baseline score, and between-group ES were computed as the ratio of the difference in mean change scores to the pooled standard deviation of change scores. A rating of "improved" on the PGI-C was used to provide an anchor-based estimate of the MIC. Data from all treatment groups were pooled for these analyses.

All study-related materials for the cognitive interviews and the collection of data used in the psychometric analysis were submitted for review by an institutional review board.

RESULTS

Cognitive Interviews

A total of 23 patients with TRD (n = 8 with experience with a fast-acting antidepressant) and 11 HCPs were interviewed. The mean age (range) of patients was 45 (20–60) years; the majority were women (65%) and white (91%); and 56% of patients were employed full-time or part-time (Table 2). The PHQ-9 scores at the day of interview indicated that patients were moderately depressed (mean PHQ-9 score = 12.4; range, 2–26).

The HCPs had diverse backgrounds and types of clinical experience, with the majority having a doctor of medicine (36%) or nursing (18%) degree. The HCPs had 14 years (mean) experience practicing in their field, mainly treating patients with depression and bipolar disorder. All HCPs had experience in using the MADRS in either general practice or clinical study settings or both, with all but 1 having more than 24 months' experience with the scale.

The majority of patients and HCPs reported that the items captured in the MADRS can fluctuate in a 24- hour

Table 3. Qualitative Interviews: Patient and HCP Reports of Variation of Depression Symptoms in a 24-Hour Period[a]

	Reported That Symptom Fluctuated in 24-Hour Period, % (n/N)			
		Patients With Prior Ketamine Experience		
Symptom	All Patients	Yes	No	All Clinicians
Apparent sadness	90 (17/19)	83 (5/6)	92 (12/13)	80 (4/5)
Reported sadness	83 (19/23)	88 (7/8)	80 (12/15)	90 (9/10)
Inner tension	95 (18/19)	100 (7/7)	92 (11/12)	100 (10/10)
Reduced sleep	53 (10/19)	63 (5/8)	45 (5/11)	55 (6/11)
Reduced appetite	78 (14/18)	100 (6/6)	67 (8/12)	64 (7/11)
Concentration difficulties	78 (14/18)	80 (4/5)	69 (9/13)	70 (7/10)
Lassitude	75 (12/16)	67 (2/3)	77 (10/13)	64 (7/11)
Inability to feel	86 (19/22)	88 (7/8)	86 (12/14)	91 (10/11)
Pessimistic thoughts	79 (15/19)	83 (5/6)	77 (10/13)	91 (10/11)
Suicidal thoughts	100 (13/13)	100 (4/4)	100 (9/9)	91 (10/11)

[a]Question to patient: Is this feeling constant, or does it come and go within a day or 24-hour period? Question to clinician: Based on your experience, does this symptom vary from day-to-day or week-to-week?
Abbreviation: HCP = health care professional.

Table 4. MADRS Descriptive Statistics at Baseline Assessment

	Study A (n = 30)		Study B (n = 65)	
MADRS Item	Mean (SD)	Median (Range)	Mean (SD)	Median (Range)
Apparent sadness	2.3 (1.45)	3 (0–4)	3.3 (1.34)	4 (0–6)
Reported sadness	2.7 (1.57)	3 (0–6)	3.5 (1.39)	4 (0–6)
Inner tension	2.3 (1.34)	3 (0–4)	2.7 (1.26)	3 (0–5)
Reduced sleep	3.0 (1.76)	4 (0–6)	2.8 (1.62)	3 (0–6)
Reduced appetite	1.1 (1.48)	0 (0–4)	1.6 (1.68)	1 (0–5)
Concentration difficulties	3.1 (1.62)	3 (0–6)	3.2 (1.51)	4 (0–6)
Lassitude	2.8 (1.51)	3 (0–5)	3.2 (1.48)	4 (0–6)
Inability to feel	2.8 (1.61)	3 (0–6)	3.5 (1.34)	4 (0–6)
Pessimistic thoughts	2.0 (1.35)	2 (0–5)	2.9 (1.32)	3 (0–5)
Suicidal thoughts[a]	0.6 (0.93)	0 (0–3)	1.0 (1.17)	1 (0–4)

[a]Patients with suicidal thoughts were not eligible for inclusion in the clinical trials.
Abbreviations: MADRS = Montgomery-Asberg Depression Rating Scale, SD = standard deviation.

period (Table 3). Specifically, over 80% of patients reported fluctuations in suicidal thoughts (100%), inner tension (95%), apparent sadness (90%), inability to feel (86%), and reported sadness (83%). The findings were similar across patients with and without experience with a fast-acting antidepressant.

In addition to assessing the variability of symptom experience, and whether a 24-hour period was appropriate to report these variations, patients were also asked about what time period would be needed to assess a clinically meaningful change in each symptom. With the exception of reduced sleep and appetite, most ($n \geq 14$, for all symptoms) patients reported that a meaningful change in depression symptoms could be assessed in a 24-hour period or less. For example, patients reported that a clinically meaningful change in reported sadness and inner tension could occur within 15 minutes, while a minimum of 1 hour was needed to determine a change in concentration difficulties, inability to feel, and pessimistic thoughts. Additionally, "a couple" of hours were needed to determine change in lassitude, and less than 24 hours to determine a change in suicidal thoughts. Similarly, HCPs also reported that meaningful changes in all symptoms could be determined in 24 hours, except reduced sleep and reduced appetite. Thus, changes in several depression symptoms are detectable by both patients and HCPs within a 24-hour time period using the MADRS-24hr.

Psychometric Analysis

The mean (SD) MADRS-24hr score at day 4 pre-dose was 22.8 (11.0) for Study A ($n = 30$) and at day 3 or day 4 pre-dose was 27.6 (8.9) for Study B ($n = 65$), corresponding to depression of moderate severity. In both studies, item-level analysis showed that MADRS-24hr scores were distributed over the possible range of levels, with little evidence of floor or ceiling effects (except suicidal thoughts, which had a mean of 0.6 [Table 4]); however, patients with suicidal thoughts were not eligible for inclusion in either study.

Overall, inter-item correlations between items on this scale were moderate. High correlations were observed between reported sadness and lassitude ($r = 0.82$; Study A), reported sadness and inability to feel ($r = 0.84$; Study A), and apparent sadness and reported sadness ($r = 0.81$; Study B). The lowest correlations were observed in Study B between inner tension and reduced sleep ($r = 0.00$), between suicidal thoughts and reduced appetite ($r = 0.05$), and between concentration difficulties and suicidal thoughts ($r = 0.05$). Results of the confirmatory factor analysis were supportive of unidimensional structure of the MADRS in both study datasets (χ^2/df) (CMIN/$df = 1.76$ and 1.95 for Studies A and B, respectively).

Internal consistency was confirmed in each dataset (Cronbach $\alpha = 0.91$ and 0.84 for Studies A and B, respectively). Due to limitations in the timing of assessments in Study B, test-retest reliability could only be assessed in Study A, in which intraclass correlation coefficients (ICC) between 2 pre-dose assessments were 0.96 and 0.91 for subpopulations defined as stable by CGI-S and PGI-S, respectively.

Construct validity was established through the observation that the MADRS-24hr total score increased monotonically with each level of depression severity, as rated by the PGI-S. The mean MADRS-24hr scores in Study A were 12.9, 24.4, and 29.2, respectively, for patients having mild, moderate, and severe depression, as rated by the PGI-S ($P = .0035$). Results were similar for patients in Study A whose severity was defined by the QIDS-SR_{14}: for the categories of none, mild, moderate, and severe/very severe depression, the mean MADRS-24hr scores were 12.1, 18.1, 27.3, and 32.8, respectively ($P < .0001$). Similar monotonic increases in mean MADRS-24hr score with increasing levels of depression severity using both the PGI-S and QIDS-SR_{14} were also observed in Study B.

Data from both studies found the MADRS-24hr to be responsive to change at the first post-dose assessment. In Study A, the within-group ES for patients who were at least "minimally improved" on CGI-I was 3.94 ($P < .0001$) compared with 1.34 ($P = .0006$) among patients defined with "no change" on the CGI-I (Table 5). The between-group ES (for improved vs no change groups) was 1.85 ($P < .0001$). Data from Study B showed a similar large, significant difference

Table 5. Responsiveness of MADRS-24hr Based on CGI-I Rating of No Change vs Improved, Study A

	No Change (n = 13)	Improved[a] (n = 17)	Between-Group ES
Baseline MADRS score, mean (SD)	32.5 (4.01)	34.4 (4.86)	
MADRS score 24 hours post-dose, mean (SD)	27.2 (5.63)	15.2 (8.97)	
Change in MADRS score, baseline to 24 hours post-dose, mean (SD)	−5.4 (4.23)	−19.2 (9.14)	1.85 ($P < .0001$)
Within-group ES, baseline to 24 hours post-dose	1.34 ($P = .0006$)	3.94 ($P < .0001$)	

[a]Definition of improved: response option of at least "minimally improved," the smallest amount of improvement in the CGI-I.
Abbreviations: CGI-I = Clinical Global Impression–Improvement, ES = effect size, MADRS = Montgomery- Asberg Depression Rating Scale, SD = standard deviation.

for within-group ES among those that improved, defined by the CGI-I (2.49 and 2.80) and between-group ES (1.45 and 2.33) for both dosing frequency arms.

Responsiveness to change was also assessed using patient groups reporting at least "improved" versus "no change" on the PGI-C, with within-group ES of 4.6 at the 24-hour post dose assessment in Study A and within-group ES > 2.5 across the 2 treatment arms in Study B.

Assessment of MIC showed a mean decrease in score ranging from 10 to 20 points on the MADRS-24hr across both trial datasets in patients who rated their condition as "improved" on the PGI-C. The corresponding MIC represented as a percentage change indicated that 30%–50% change in MADRS-24hr score could be interpreted as meaningful improvement.

DISCUSSION

This study provides data supporting the appropriateness of the MADRS-24hr scale to assess change in MDD symptomatology over a 24-hour recall period. Results show that depressive symptoms can fluctuate during a 24-hour period and that this window could be used to capture meaningful changes in the majority of MDD symptoms assessed within the MADRS. Although the MADRS-24hr can be a useful tool for capturing rapid improvements of symptoms, follow-up over several days to weeks may be necessary to confirm that this change is real and can be sustained. Patients and clinicians indicated that 1 to 2 weeks was an acceptable period to capture sustained change in sadness, reduced sleep, and lassitude, but pessimistic with suicidal thoughts required a longer time period. The intent of this work was not to propose modification of the items contained within the MADRS since the full spectrum of symptoms is still important to evaluate treatment effects in the longer term; rather, we have shown that this instrument with a shortened recall period can be used to assess rapid changes in depressive symptomatology.

Assessment of the psychometric properties of the MADRS-24hr showed that this instrument had acceptable reliability and validity and was responsive to change following an intervention, further supporting the feasibility of evaluation over a 24-hour period and measurement using this recall period in assessment of treatment effects.

This study adds to the prior research that involved work using the MADRS with a shortened recall period in a clinical trial setting[15] and provides further evidence to support the use of the MADRS with a shortened recall period to assess change in depressive symptoms in patients with bipolar II disorder.[23] Although this study found that the MADRS-24hr had acceptable psychometric properties, the design of the clinical trials and small sample size limited the ability to fully assess the psychometric properties (eg, test-retest reliability could not be assessed using data from Study B). An additional limitation of the study was the limited number of patients and HCPs who participated in the cognitive interviews. Although the sample was adequate to confirm content validity of the instrument, replication of these findings in a larger sample would add to this body of research. Given these limitations, additional studies are warranted, including studies of more geographically diverse populations of patients with MDD and across broader sociodemographic characteristics.

Overall, these findings suggest that a 24-hour recall version of the MADRS can be used to detect rapid treatment change following administration of a drug with a fast onset of action.

Submitted: July 21, 2015; accepted February 11, 2016.

Drug names: ketamine (Ketalar and others), quetiapine (Seroquel and others).

Author contributions: **Dr Johnson** contributed to the study design and data interpretation, wrote the first draft of the manuscript, and reviewed the final manuscript. **Ms Jamieson** participated in study design, data analysis and interpretation, and writing and review of the manuscript. **Ms Howard** was responsible for the content validity study concept and design and played an integral part in the analysis and interpretation of the content validity results. In addition, she contributed to the writing and review of the manuscript. **Mr Devine** participated in the study design; played a primary role in the collection, analysis, and interpretation of the content validation data; and contributed to the writing and review of the manuscript. **Dr Ho** was responsible for the design of and conducted the psychometric data analyses. **Mr Saretsky** participated in the study design, data collection, analysis, interpretation, and writing and review of the manuscript. All authors meet ICMJE criteria, and all those who fulfilled those criteria are listed as authors. All authors had access to the study data and made the final decision about where to present these data.

Potential conflicts of interest: **Dr Johnson** was an employee of Janssen Global Services during the conduct of this study. **Ms Jamieson** is an employee of Janssen Global Services and owns stocks in the company. **Mr Devine**, **Ms Howard**, and **Mr Saretsky** are employees of ICON Clinical Research (formerly Oxford Outcomes) and received payment from Janssen Global Services for providing research services as an employee of ICON. **Dr Ho** is an independent statistical consultant who received payment from Janssen Global Services for his work on the statistical analyses.

Funding/support: Janssen Research & Development was the sponsor for this study.

Role of the sponsor: Janssen Research & Development had a role in the design and conduct of the study; in the collection, analysis, and interpretation of the data; and in the preparation, review, or approval and in the decision to submit the manuscript for publication.

Acknowledgments: The authors acknowledge Pravin Bolshete, BAMS, and Shalini Nair, PhD (SIRO Clinpharm Pvt. Ltd.) for writing assistance and Wendy Battisti, PhD, and Ellen Baum, PhD (Janssen Research & Development, LLC) for additional editorial support for the development of this manuscript. Writing support was funded by Janssen Research & Development, LLC.

REFERENCES

1. Andrade L, Caraveo-Anduaga JJ, Berglund P, et al. The epidemiology of major depressive episodes: results from the International Consortium of Psychiatric Epidemiology (ICPE) Surveys. *Int J Methods Psychiatr Res*. 2003;12(1):3–21.
2. Kessler RC, Berglund P, Demler O, et al; National Comorbidity Survey Replication. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA*. 2003;289(23):3095–3105.
3. Kubitz N, Mehra M, Potluri RC, et al. Characterization of treatment resistant depression episodes in a cohort of patients from a US commercial claims database. *PLoS One*. 2013;8(10):e76882.
4. Corey-Lisle PK, Birnbaum HG, Greenberg PE, et al. Identification of a claims data "signature" and economic consequences for treatment-resistant depression. *J Clin Psychiatry*. 2002;63(8):717–726.
5. Crown WH, Finkelstein S, Berndt ER, et al. The impact of treatment-resistant depression on health care utilization and costs. *J Clin Psychiatry*. 2002;63(11):963–971.
6. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry*. 1979;134:382–389.
7. Galinowski A, Lehert P. Structural validity of MADRS during antidepressant treatment. *Int Clin Psychopharmacol*. 1995;10(3):157–161.
8. Khan A, Khan SR, Shankles EB, et al. Relative sensitivity of the Montgomery-Asberg Depression Rating Scale, the Hamilton Depression Rating Scale and the Clinical Global Impressions rating scale in antidepressant clinical trials. *Int Clin Psychopharmacol*. 2002;17(6):281–285.
9. Bondolfi G, Jermann F, Rouget BW, et al. Self- and clinician-rated Montgomery-Asberg Depression Rating Scale: evaluation in clinical practice. *J Affect Disord*. 2010;121(3):268–272.
10. Svanborg P, Asberg M. A comparison between the Beck Depression Inventory (BDI) and the self-rating version of the Montgomery Asberg Depression Rating Scale (MADRS). *J Affect Disord*. 2001;64(2–3):203–216.
11. Hedman E, Ljótsson B, Blom K, et al. Telephone versus internet administration of self-report measures of social anxiety, depressive symptoms, and insomnia: psychometric evaluation of a method to reduce the impact of missing data. *J Med Internet Res*. 2013;15(10):e229.
12. Holländare F, Andersson G, Engström I. A comparison of psychometric properties between internet and paper versions of two depression instruments (BDI-II and MADRS-S) administered to clinic patients. *J Med Internet Res*. 2010;12(5):e49.
13. Mundt JC, Katzelnick DJ, Kennedy SH, et al. Validation of an IVRS version of the MADRS. *J Psychiatr Res*. 2006;40(3):243–246.
14. Silverstone PH. Depression increases mortality and morbidity in acute life-threatening medical illness. *J Psychosom Res*. 1990;34(6):651–657.
15. Suppes T, Ketter TA, Gwizdowski IS, et al. First controlled treatment trial of bipolar II hypomania with mixed symptoms: quetiapine versus placebo. *J Affect Disord*. 2013;150(1):37–43.
16. Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. US Food & Drug Administration Web site. http://webcache.googleusercontent.com/search?q=cache:vGr1SNlJsLwJ:www.fda.gov/downloads/Drugs/.../Guidances/UCM193282.pdf+&cd=1&hl=en&ct=clnk&gl=in. December 2009. Accessed January 29, 2016.
17. Singh JB, Fedgchin M, Daly E, et al. Intravenous esketamine in adult treatment-resistant depression: a double-blind, double-randomization, placebo-controlled study. *Biol Psychiatry*. 2016;80(6):424–431.
18. 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.
19. Joffe H, Yerdley L. Content and thematic analysis. In: Marks D, Yardley L, eds. *Research Methods for Clinical and Health Psychology*. London, UK: Sage; 2004:56–68.
20. Singh JB, Fedgchin M, Daly EJ, et al. A double-blind, randomized, placebo-controlled, dose-frequency study of intravenous ketamine in patients with treatment-resistant depression. *Am J Psychiatry*. 2016;173(8):816–826.
21. Guy W. *ECDEU Assessment Manual for Psychopharmacology*. Revised Edition. Washington, DC: US Department of Health, Education, and Welfare; 1976.
22. Wyrwich KW, Bullinger M, Aaronson N, et al; Clinical Significance Consensus Meeting Group. Estimating clinically significant differences in quality of life outcomes. *Qual Life Res*. 2005;14(2):285–295.
23. Revicki DA, Erickson PA, Sloan JA, et al; Mayo/FDA Patient-Reported Outcomes Consensus Meeting Group. Interpreting and reporting results based on patient-reported outcomes. *Value Health*. 2007;10(suppl 2):S116–S124.
24. Rush AJ, Trivedi MH, Ibrahim HM, et al. The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. *Biol Psychiatry*. 2003;54(5):573–583.
25. Trivedi MH, Rush AJ, Ibrahim HM, et al. The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR), and the Quick Inventory of Depressive Symptomatology, Clinician Rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. *Psychol Med*. 2004;34(1):73–82.