Published in final edited form as:
J Affect Disord. 2006 October ; 95(1-3): 115–118.

Making Clinicians Lives Easier: Guidance on Use of the QIDS Self-Report in Place of the MADRS

Thomas J. Carmody, Ph.D.[a,*], **A. John Rush, M.D.**[a], **Ira H. Bernstein, Ph.D.**[b], **Stephen Brannan, M.D.**[c], **Mustafa M. Husain, M.D.**[a], and **Madhukar H. Trivedi, M.D.**[a]

aDepartment of Psychiatry, University of Texas Southwestern Medical Center, Dallas, TX*

bUniversity of Texas at Arlington*

cCyberonics, Inc.*

Abstract

Background—The ability to convert total scores from one scale to another facilitates the interpretation of research findings and facilitates the use of systematic measurement in clinical practice.

Methods—Item Response Theory methods were used to convert total scores between the 16-item Quick Inventory of Depressive Symptomatology (QIDS-SR_{16}) and the Montgomery Asberg Depression Rating Scale (MADRS) total scores. Data were obtained from a sample of 233 outpatients with highly treatment-resistant, nonpsychotic major depressive episodes participating in a one-year open label study of vagus nerve stimulation to augment psychotropic medication treatment.

Results—MADRS total scores averaged 31.9 (SD=6.7) at baseline and 21.9 (SD=11.0) at one year. QIDS-SR_{16} total scores averaged 17.6 (SD=3.6) at baseline and 12.5 (SD=5.8) at one year. Based on one-year data (or exit if the patient did not complete one year), corresponding QIDS-SR_{16} and MADRS total scores were presented for each possible QIDS-SR_{16} and MADRS total score. A QIDS-SR_{16} total score of 5 was comparable to a MADRS total score of 7 or 8 (7.5).

Limitation—The degree to which these results generalize to less treatment-resistant samples is unknown.

Conclusion—The conversion of QIDS-SR_{16} and MADRS total scores provides a basis for clinicians who wish to use the QIDS-SR_{16} to understand what MADRS total scores reported in clinical trials approximate QIDS-SR_{16} total scores obtained with their patients.

Keywords

Montgomery-Äsberg Depression Rating Scale (MADRS); 16-item Quick Inventory of Depressive Symptomatology - Self-report (QIDS-SR16); item response theory; classical test theory; psychometrics; total score conversion

INTRODUCTION

Several depression rating scales are presently in use both in clinical research and in the management of patients with depression. The interpretation of research findings and individual patient level assessments based on these rating scales would be facilitated by the ability to convert total scores on one scale to total scores on other scales. For example, published findings

Address correspondence to: Thomas J. Carmody, Ph.D., Department of Psychiatry, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX 75390-9086, Tel: (214) 648-4610, Fax: (214) 648-4612, E-mail: thomas.carmody@utsouthwestern.edu

with a clinician rated scale could be converted into results based on a patient self-rated scale. Also, threshold total scores for remission, and mild, moderate, and severe depression for one scale could be identified by reference to corresponding thresholds for another scale.

This report uses item response theory (IRT) methods (Orlando et al., 2000) on a sample of 233 treatment-resistant depressed outpatients to equate a relatively new, but increasingly used, brief 16-item self-report — the Quick Inventory of Depressive Symptomatology–Self-Report ($QIDS\text{-}SR_{16}$) (Rush et al., 2003;Trivedi et al., 2004) — to a more widely used clinician rating scale, the Montgomery Asberg Depression Rating Scale (MADRS) (Montgomery and Asberg, 1979).

METHODS

Study

Data were obtained from a study of vagus nerve stimulation as a treatment for depression used in addition to ongoing medication regimens (Rush et al., 2005b). Data were obtained at one year (or the date closest to one year) following study initiation. At one year the study was open label with raters unblinded to treatment. Patients with no post-baseline data were excluded. Study participants were adult outpatients (18–75 years old) with highly treatment-resistant, nonpsychotic major depressive episodes (MDEs). Diagnoses of depression were determined using the Structured Clinical Interview for DSM-IV (SCID) (First et al., 1994). Data were supplied by Cyberonics, Inc. and analyzed by the first author (TC).

Scales

The 16-item Quick Inventory of Depressive Symptomatology (QIDS) assesses the nine DSM-IV criteria symptom domains: sad mood, concentration, self-outlook, suicidal ideation, involvement, energy/fatigability, sleep disturbance (4 items: initial, middle, late insomnia, and hypersomnia), appetite/weight increase/decrease (4 items), and psychomotor agitation/retardation (2 items) (Rush et al., 2003;Trivedi et al., 2004). Both self-report ($QIDS\text{-}SR_{16}$) and clinician versions ($QIDS\text{-}C_{16}$) are available using identical items (www.ids-qids.org). The psychometric properties of these brief scales have been extensively evaluated (Rush et al., 2003;Trivedi et al., 2004;Rush et al., 2005a) and IRT analyses have reported tables by which to convert total scores on the QIDS to total scores on several versions of the Hamilton Rating Scale for Depression (HRSD) (Hamilton 1960,1967). The 10-item clinician rated Montgomery Asberg Depression Rating Scale (MADRS) assesses most but not all of the DSM-IV criterion symptoms (Montgomery and Asberg, 1979). The MADRS has very good psychometric properties (Khan et al., 2002;Galinowski and Lehert, 1995) and is widely used in clinical trials both in Europe and the United States.

Statistical Methods

Samejima's graded IRT model (Samejima, 1997) item parameters were estimated for each item of the $QIDS\text{-}SR_{16}$ and MADRS and then used to generate an IRT score for each possible total score on each measure according to the procedure of Orlando et al. (2000) (and associated software). The IRT score, usually called theta, is a unitless measure of depression commonly scaled to have mean 0 and a standard deviation of 1. Finally, total scores were equated by matching the corresponding IRT scores. When an exact match between IRT scores was not available, best judgment was used to equate the scales taking into account the matching of total scores immediately above and below the total score in question.

The unidimensionality assumption of the IRT approach was assessed using parallel analysis (Humphreys and Montanelli, 1975) to infer the number of factors. This is a more recently developed alternative to the traditional eigenvalue greater than one criterion. In parallel

analysis, instead of choosing factors with eigenvalues greater than one, factors are chosen with eigenvalues greater than would be expected to arise by change alone. Specifically, principal component eigenvalues from the real data are compared to eigenvalues from simulated datasets with the same number of observations and items as the real data and where correlations between all items are expected to be zero. A total of 1000 such simulated datasets were generated. The dimensionality is defined as the number of principal components whose real data eigenvalues exceed the average of the simulated data eigenvalues.

RESULTS

Sample Characteristics

Most patients in this study were diagnosed with nonpsychotic major depressive disorder: 208/233 (89.3%) patients. However, 25/233 (10.7%) were in a depressed phase of bipolar I (n=12) or bipolar II (n=13) disorder. Altogether, 62.2% (145/233) were female and 96.6% were Caucasian with an average age of 47.2 years (SD=8.9) (range: 24 to 72). Length of illness averaged 25.0 years (SD=12.0) with an average current episode of 3.8 years (SD=4.0). These patients were highly treatment-resistant, having sustained between two and six trials of known effective treatments delivered at adequate doses and durations in the current MDE as assessed by the Antidepressant Treatment History Form (Sackeim, 2001). Patients had received over 12 different medications on average in the current MDE when all clinical potential antidepressant treatments were counted (Rush et al, 2005b).

$HRSD_{17}$ total scores averaged 21.9 (SD=4.4) (range: 13 to 37) at baseline and 15.6 (SD=7.1) (range: 2 to 33) at one year. MADRS total scores averaged 31.9 (SD=6.7) (range: 14 to 50) at baseline and 21.9 (SD=11.0) (range: 0 to 47) at one year. QIDS-SR_{16} total scores averaged 17.6 (SD=3.6) (range: 8 to 27) at baseline and 12.5 (SD=5.8) (range: 2 to 26) at one year.

Dimensionality

For the MADRS, the first real data eigenvalue of 5.73 was much larger than the first simulated data eigenvalue of 1.34, while the second real data eigenvalue of 1.06 was smaller than the second simulated data eigenvalue of 1.23. Therefore, the MADRS was determined to be unifactorial. Comparison of simulated versus real data eigenvalues also showed the QIDS-SR_{16} to be unifactorial. The first two eigenvalues were 1.31 and 1.21 (simulated) versus 4.77 and 1.00 (real).

Conversion Table

Table 1 summarizes the IRT conversions for QIDS-SR_{16} and MADRS total scores. A QIDS-SR_{16} total score of 5 (remission threshold) was comparable to a MADRS total score of 7 or 8 (7.5). QIDS-SR_{16} depression severity thresholds have been suggested (www.ids-qids.org) of 6 to10 for mild, 11 to 15 for moderate, 16 to 20 for severe, and 21+ for very severe depression. Using Table 1, the corresponding MADRS thresholds would be 9 to 18 for mild, 19 to 27 for moderate, 28 to 36 for severe, and 37+ for very severe depression.

CONCLUSION

QIDS-SR_{16} and MADRS total scores were equated using a sample of 233 treatment-resistant, nonpsychotic depressed outpatients, providing a basis for clinicians who wish to use the QIDS-SR_{16} to understand what MADRS total scores reported in clinical trials approximate QIDS-SR_{16} total scores obtained with their patients. Whether these results generalize to less treatment-resistant samples deserves study.

Acknowledgements

The authors wish to express appreciation to Cyberonix, Inc. for providing the data and to Fast Word, Inc. Dallas, TX for secretarial support. These analyses were supported in part by the National Institute of Mental Health (NIMH), National Institutes of Health (MH-68851 to the University of Texas Southwestern Medical Center at Dallas, A. John Rush, M.D., PI, and by MH-68852 to the University of Texas at Arlington, Ira H. Bernstein, Ph.D., PI).

The authors personally conducted these analyses and received no support for conducting these analyses or preparing this report. A. John Rush, M.D. has received payments as a consultant and speaker for Cyberonics Inc. Madhukar Trivedi, M.D. has received payments as a consultant to Cyberonics. Steve Brannan, M.D. is an employee and stockholder of Cyberonics, Inc.

References

First, MB.; Spitzer, R.; Gibbon, M.; Williams, J. Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) Patient. NY State Psychiatric Institute Biometrics Research Department; New York, NY: 1994.

Galinowski A, Lehert P. Structural validity of MADRS during antidepressant treatment. Int Clin Psychopharmacol 1995;10:157–161. [PubMed: 8675968]

Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry 1960;23:56–62. [PubMed: 14399272]

Hamilton M. Development of a rating scale for primary depressive illness. Br J Soc Clin Psychol 1967;6 (4):278–296. [PubMed: 6080235]

http://www.ids-qids.org/Severity_Thresholds.pdf. Retrieved February 16, 2006.

Humphreys LG, Montanelli RG Jr. An investigation of the parallel analysis criterion for determining the number of common factors. Multivariate Behav Res 1975;10:193–206.

Khan A, Khan SR, Shankles EB, Polissar NL. 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;17:281–285. [PubMed: 12409681]

Montgomery SA, Äsberg M. A new depression scale designed to be sensitive to change. Br J Psychiatry 1979;134:382–389. [PubMed: 444788]

Orlando M, Sherbourne CD, Thissen D. Summed-score linking using item response theory: application to depression measurement. Psychol Assess 2000;12:354–359. [PubMed: 11021160]

Rush AJ, Bernstein IH, Trivedi MH, Carmody TJ, Wisniewski S, Mundt JC, Shores-Wilson K, Biggs MM, Nierenberg AA, Fava M. An evaluation of the Quick Inventory of Depressive Symptomatology and the Hamilton Rating Scale for Depression: a STAR*D report. Biol Psychiatry. 2005aEpub ahead of print 2005 Sep 28

Rush AJ, Sackeim HA, Marangell LB, George MS, Brannan SK, Davis S, Lavori P, Howland R, Kling MA, Rittberg B, Carpenter L, Ninan P, Moreno F, Schwartz T, Conway C, Burke M, Barry J. Effects of 12 months of vagus nerve stimulation in treatment-resistant depression: a naturalistic study. Biol Psychiatry 2005b;58(5):355–363. [PubMed: 16139581]

Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, Markowitz JC, Ninan PT, Kornstein S, Manber R, Thase ME, Kocsis JH, Keller MB. 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. [PubMed: 12946886]

Sackeim HA. The definition and meaning of treatment-resistant depression. J Clin Psychiatry 2001;62 (Suppl 16):10–17. [PubMed: 11480879]

Samejima, F. Graded response model. In: van Linden, W.; Hambleton, RK., editors. Handbook of Modern Item Response Theory. Springer-Verlag; New York, NY: 1997. p. 85-100.

Trivedi MH, Rush AJ, Ibrahim HM, Carmody TJ, Biggs MM, Suppes T, Crismon ML, Shores-Wilson K, Toprac MG, Dennehy EB, Witte B, Kashner TM. 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. [PubMed: 14971628]

Table 1

IRT Conversion between the QIDS-SR_{16} and MADRS using 1-year data (n=233)

Conversion QIDS-SR_{16} – MADRS	
0	0
1	1
2	2
3	3 or 4
4	5 or 6
5	7 or 8
6	9 or 10
7	11 or 12
8	13 or 14
9	15 or 16
10	17 or 18
11	19 or 20
12	21
13	22 or 23
14	24 or 25
15	26 or 27
16	28 or 29
17	30 or 31
18	32 or 33
19	34
20	35 or 36
21	37 or 38
22	39 or 40
23	41 or 42
24	43 or 44
25	45 or 46
26	47 or 48
27	49 to 60