

# Functional Outcomes in MDD: Established and Emerging Assessment Tools

Jean Endicott, PhD; and Kathleen M. Dorries, PhD

## Abstract

Currently, the primary efficacy measures for antidepressant clinical trials predominantly assess changes in mood symptoms in patients with major depressive disorder (MDD). When considering treatment options, however, patients and clinicians value improvement in function in important life domains as highly as symptom reduction. MDD patients report that they consider a return to normal functioning an important indicator of remission from depressive episodes. Indeed, many researchers now regard assessment of both mood symptoms and functional outcomes essential to measuring treatment-related improvement and remission from MDD. However, function is a very broad concept. Investigators must consider multiple issues in designing or selecting an instrument that measures function adequately and appropriately in their particular study population. The assessment tool should include dimensions of functioning that are relevant and likely to improve with the treatment, and the instrument should have demonstrated reliability and validity, good discrimination among patients, and sensitivity to meaningful improvement in functioning. The inclusion of well-chosen functional outcome measures in clinical trials will improve the assessment of impairment and improvement with treatment, and provide patients and clinicians with important information about the efficacy of antidepressant treatments.

(*Am J Manag Care. 2009;15:S328-S334*)

For author information and disclosures, see end of text.

## Introduction

The *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR)*<sup>1</sup> diagnostic criteria for major depressive disorder (MDD) include the presence of a number of symptoms that are severe enough to markedly impair social, occupational, or other functioning. There are as yet no physical markers that can be identified or measured in the diagnosis of MDD, or in the assessment of efficacy of drugs used to treat the disease, and outcome measures for antidepressant clinical trials therefore utilize instruments designed to measure change in psychological markers for the disease. Generally, the primary outcome measures for antidepressant clinical trials reflect changes in mood symptoms rather than functional deficits associated with those symptoms. The standard instrument most often used for the primary outcome measure in MDD clinical trials is the 17-item Hamilton Rating Scale for Depression (HAM-D<sub>17</sub>),<sup>2</sup> with the Montgomery-Åsberg Depression Rating Scale (MADRS)<sup>3</sup> and Clinical Global Impressions-Improvement (CGI-I)<sup>4</sup> often used as key secondary outcomes. Clinical trials often include instruments that measure functional outcomes as secondary measures, but results based on those instruments are not generally considered by the US Food and Drug Administration or its equivalent in other countries in drawing conclusions about the efficacy of treatment. Research investigating the relationships between functional outcomes and symptom reduction, remission, and recurrence of MDD suggest, however, that the current focus on the mood symptoms of the *DSM-IV-TR* diagnostic criteria, but not their impact on functioning, is too narrow a look at the pathology of the disease.<sup>5-9</sup> A definition of remission based on the HAM-D<sub>17</sub> (HAM-D<sub>17</sub> score  $\leq 7$ ) was proposed to address the need for a standard criterion for when a patient “no longer meets syndromal criteria” for MDD, indicating that “no increase in the intensity of the treatment regimen is required.”<sup>10</sup> Remission rate, based on this definition (or other corresponding definitions for other scales<sup>11</sup>), is now a common outcome measure in clinical trials. Measurement of mood symptoms alone, however, can result in failure to detect significant impairment.

Depressive symptom reduction and improvement in function were found to be moderately correlated in a meta-analysis of studies that reported correlations between depression symptom measures and functional outcome measures, suggesting that functional measures are not redundant with measures of symptom reduction, but

instead capture independent information.<sup>9</sup> For example, correlations between scores on the Endicott Work Productivity Scale and HAM-D<sub>17</sub> scores were 0.46 and 0.57 for MDD patients after 12 weeks of cognitive behavioral therapy or nefazodone, respectively ( $r = 0.48$  for combined treatment).<sup>12</sup> Correlations between HAM-D<sub>17</sub> scores and scores on the Social Adjustment Scale (SAS)<sup>13</sup> (assessing function in the work, social, and family domains) were 0.577 for MDD patients after 6 weeks of treatment with fluoxetine or nortriptyline.<sup>14</sup> Judd and colleagues<sup>15</sup> found that the correlation between global functioning measures and depressive symptom severity was high ( $r = 0.72$ ), but correlations were lower for specific areas of functioning ( $r$ 's not reported). In some patient populations, mood symptoms may provide less information than functional impairment about the severity of depression: Gallo and colleagues<sup>16</sup> found in a large ( $N = 1612$ ), prospective study that participants aged 50 years or older with symptoms of MDD but no sadness or dysphoria had an increased risk of functional impairment over a 13-year follow-up period, consistent with the findings of the National Institutes of Health Consensus Panel that elderly MDD patients are less likely compared with younger adults to present with dysphoria.<sup>17</sup> Judd and colleagues found that MDD patients ( $N = 371$ ) classified as remitters with minimal residual depressive symptoms had notable impairment in work functioning based on the Longitudinal Interval Follow-up Evaluation (LIFE).<sup>15</sup>

While measures of functioning are only moderately correlated with symptom reduction, they may be important predictors of a patient's course of depression, including both recovery from, and possible recurrence of, depressive episodes. In an observational study of 231 patients with MDD, greater impairment in psychosocial functioning, as measured by the LIFE-Range of Impaired Functioning Tool (LIFE-RIFT), was associated with a lower likelihood of remission or recovery from an episode of depression over the assessment period of 6 months to 1 year.<sup>18</sup> In a separate analysis of data from the above-mentioned study, 290 patients in recovery from an episode of MDD (8 consecutive weeks with  $\leq 2$  mild symptoms of MDD) were administered the LIFE-RIFT at 6-month intervals for 5 years, then yearly for 10 years.<sup>19</sup> Results showed that residual functional impairment significantly increased the risk of recurrence within 6 to 12 months.

Measuring functional outcomes also allows the assessment of the meaningful impact of MDD on overall well-being. MDD patients report that they would consider a return to normal functioning an important indicator of remission. Zimmerman and colleagues<sup>20</sup> designed a checklist of potential factors indicating remission from MDD by asking patients

how they knew, or would know, if their depression was in remission. Five of the 16 factors suggested by the patients were related to improvements in function. The investigators then asked 535 MDD outpatients to rate the 16 items from 0 (not very important) to 2 (very important) for determining whether someone is in remission from depression. The factors rated "very important" by the largest percentages of patients were as follows: (1) "presence of positive mental health (eg, optimism, vigor, self-confidence)" (77.3%), (2) "feeling like your usual, normal self" (75.6%), and (3) "return to usual level of functioning at work, home, or school" (74.3%), demonstrating that patients themselves consider normal functioning critical to their own understanding of remission.<sup>21</sup>

### What Is Function?

Function is a very broad concept, and it is used by different investigators to cover different types of domains. Areas commonly included in the concept of "function" are role functioning (the ability to carry out tasks and activities associated with the patient's role as, for example, a worker, parent, student, etc), personal care (eg, bathe, shave, get dressed), social functioning (family and social relationships), and satisfaction with one's performance in a particular life area (also related to life satisfaction or quality of life). A useful definition of the term *function* is: to perform in a required or expected manner, where "perform" essentially means to do a task or to carry something out. Areas of function that comprise possible outcome measures include social functioning, functioning within family relationships, work (or school) functioning, self-care (for a chronic population), and cognitive functioning.

### Assessment Tools

There are several well-established instruments (Table<sup>1,13,22-32</sup>) that are commonly used for assessing function in MDD clinical trials: the Sheehan Disability Scale (SDS),<sup>22</sup> the SAS,<sup>13,23</sup> the Global Assessment of Functioning (GAF) scale,<sup>24</sup> and the Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36).<sup>25</sup> Although not all of these scales were designed to measure function only—the SF-36, for example, includes pain, self-reported health, and psychological symptom subscales—they are used in MDD trials as functional end points. The most common functional outcomes scale in psychopharmacology research, the SDS, was developed as a very brief (3-item) but sensitive measure of the effect of psychiatric or other chronic illness on 3 areas of function: work/studies, social life, and family life/home responsibilities.<sup>22,33</sup> Patients rate how disruptive their symptoms have been in each of these domains using discretized visual analog scales, with ordered

■ **Table.** Selected Assessment Tools for Measuring Functional Outcomes

Instrument	No. of Items	Domains Assessed
Social and Occupational Functioning Assessment Scale <sup>1</sup>	Single scale	Social and occupation functioning
Sheehan Disability Scale <sup>22</sup>	3	Work/studies Social life Family life/home responsibilities
Social Adjustment Scale <sup>13,23</sup>	48	Work (paid, homemaker, student) Social and leisure activities Relationship with extended family Marital role Parental role Role as a member of a family unit
Global Assessment of Functioning scale <sup>24</sup>	Single scale	Combined severity of symptoms and impaired functioning
Medical Outcomes Study 36-Item Short-Form Health Survey <sup>25</sup>	36	Physical functioning Physical role functioning Bodily pain General health Vitality Social functioning Emotional role functioning Mental health
World Health Organization Disability Assessment Schedule II <sup>26</sup>	36	Cognition Mobility Self-care Interpersonal interactions Life activities Participation in society
Endicott Work Productivity Scale <sup>27</sup>	25	Work efficiency and productivity
Longitudinal Interval Follow-up Evaluation–Range of Impaired Functioning Tool <sup>28</sup>	9	Work Interpersonal relationships Overall satisfaction with functioning Recreation
Life Functioning Questionnaire <sup>29</sup>	14	Duties at work/school Duties at home Leisure time with family Leisure time with friends
Health of the Nation Outcome Scales <sup>30</sup>	12	Behavior Impairment Symptoms Social functioning
UCLA Social Attainment Survey <sup>31</sup>	7	Peer relationships Romantic relationships Involvement with activities
Functioning Assessment Short Test <sup>32</sup>	24	Autonomy Occupational functioning Cognitive functioning Financial issues Interpersonal relationships Leisure time

UCLA indicates University of California, Los Angeles.

verbal descriptors ranging from “not at all” (0) to “extremely” (10). Total scores are calculated for patients who complete all 3 items. The SDS has been used widely in clinical trials of MDD and anxiety disorders.<sup>33</sup>

The SAS and the SAS self-report version (SAS-SR) both include 48 items in 6 domains: work (paid, homemaker, student); social and leisure activities; relationship with extended family; marital role; parent role; and role as

a member of a family unit.<sup>23</sup> Items are rated on a 5-point, Likert-type scale (higher values indicating poorer function), and means are calculated for each domain and for an overall score. A short version and a screening version of the SAS-SR are available.<sup>13,34</sup>

The GAF consists of a single 100-point scale combining severity of symptoms and impaired functioning.<sup>24</sup> The patient's level of functioning is rated from a hypothetical sickest possible (1) to healthiest possible (100). The scale is divided into 10-point intervals (1-10, 11-20, etc), and each interval is labeled with defining characteristics. More than the SDS and the SAS, the GAF combines both severity of symptoms and degree of impaired functioning in a single score. A patient-rated version of the GAF shows good agreement with the clinician-rated scale.<sup>35</sup>

The SF-36 is a self-report tool that assesses 8 domains—physical functioning, physical role functioning, bodily pain, general health, vitality, social functioning, emotional role functioning, and mental health—with 2 to 10 items per domain. Different items are scored according to different scales, ranging from 2 (yes or no) to 6 levels.<sup>25</sup> This scale includes a mixture of satisfaction, functioning, and symptoms, and is commonly considered by investigators to be a quality-of-life measure.<sup>7,36,37</sup> However, the validated 12-item version of the SF-36 (which includes a subset of the 36 items but reproduces the average summary scores<sup>38</sup>) was compared with the Quality of Life Enjoyment and Satisfaction Questionnaire (Q-LES-Q), developed specifically to assess self-reported quality of life,<sup>39</sup> in patients with MDD in the Sequenced Treatment Alternatives to Relieve Depression trial.<sup>40</sup> The investigators reported a lack of correlation between items on the 2 instruments, which they concluded was due to a difference in the focus of the 2 instruments: the Q-LES-Q measures enjoyment and satisfaction and the Short Form Health Survey measures self-reported functioning.

Numerous other instruments have been designed to assess functioning in clinical trials or practice. Among those that are less frequently used, but might be considered emerging instruments, are the Social and Occupational Functioning Assessment Scale of the *DSM-IV-TR*, a single scale (0-100) similar in structure to the GAF, assessing social and occupation functioning<sup>1</sup>; the World Health Organization Disability Assessment Schedule II (WHODAS II), which includes 36 items within 6 domains (cognition, mobility, self-care, interpersonal interactions, life activities [domestic responsibilities, leisure, work], and participation in society)<sup>26,41</sup>; the Endicott Work Productivity Scale, a measure of “presenteeism,” or productivity lost while present at work,<sup>9</sup> using 25 five-point items covering behavior and attitudes that are likely to

impair productivity and efficiency on work activities<sup>27</sup>; the LIFE-RIFT, a semistructured interview that includes 9 items assessing 4 domains (work [employment, household activities, school], interpersonal relationships [spouse, children, other relatives, friends], overall satisfaction with functioning, and recreation)<sup>28</sup>; the Life Functioning Questionnaire (LFQ), with 14 items in 4 domains (duties at work/school, duties at home, leisure time with family, leisure time with friends)<sup>29</sup>; the Health of the Nation Outcome Scales (HoNOS), which consists of 12 five-point (0-4) scales that cover 4 domains (behavior, impairment, symptoms, and social functioning)<sup>30</sup>; the University of California Los Angeles Social Attainment Survey, which includes 7 five-point (1-5) items that yield 3 domain scores (peer relationships, romantic relationships, involvement with activities) and a total score<sup>31,42</sup>; and the Functioning Assessment Short Test (FAST), a 24-item instrument covering 6 domains (autonomy, occupational functioning, cognitive functioning, financial issues, interpersonal relationships, and leisure time).<sup>32</sup>

### Issues in Measuring Function

The wide variety of instruments used to measure functional impairment, as well as the numerous differences between them, highlight several issues that arise in selecting a tool for measuring function in research or in the clinic. First, different functions are applicable to different patient populations, based on social/occupational expectations or abilities. For example, one of the limitations of the SAS is that it shows sex bias; in a validation study for the LFQ, men were significantly less likely to complete the “Work at Home” section of the SAS-SR compared with the “Duties At Home” section of the LFQ.<sup>29</sup> The LFQ includes a more balanced set of items in that domain, whereas the SAS-SR includes mainly tasks traditionally associated with women. In addition, the SAS-SR instructs respondents to choose between “worker for pay,” “housewife,” “student,” “retired,” and “unemployed,” to describe their principal role, whereas the LFQ allows multiple descriptors. In one sample, over one third of the women completed both “worker” and “housewife” sections in the SAS despite instructions to choose just one.<sup>29</sup> Instruments may also be more or less appropriate for assessing functional impairment in elderly patients compared with mixed-aged populations. In a survey of community-dwelling elderly compared with a normative mixed-age adult sample, the older adults showed greater impairment in overall social functioning on the SAS-SR, but no difference between the groups was seen after 3 items on dating and sexual activity were excluded from the analysis.<sup>34</sup> Because cognitive symptoms are particularly likely in elderly MDD patients,<sup>1</sup>

## Reports

instruments that include a cognitive domain, such as the WHODAS II or FAST, can be useful for assessing function in the elderly.<sup>26,32,43</sup> For some scales, different versions have been developed for specific populations (eg, HoNOS 65+ and HoNOS for Children and Adolescents).<sup>44,45</sup> A more general scale, however, would be more useful for studying diverse patient populations or for making comparisons across studies through inclusion in a meta-analysis.

A second issue in measuring function is the distinction between functioning and quality-of-life or satisfaction measures. Functioning (particularly social functioning) and satisfaction are often considered components of health-related quality of life: poor social functioning and low satisfaction adversely affect a patient's quality of life.<sup>5,7,46</sup> However, functioning and satisfaction can vary independently; for example, a patient may have measurable impairment in many activities, and still be quite satisfied with his or her situation in life. Furthermore, patient satisfaction is subjective and therefore patient-rated, whereas functioning can be more objectively measured.<sup>5,7,46</sup> There are differences in the utility or appropriateness of those items that capture function using objective measures, such as time lost from work, versus those that capture satisfaction using subjective measures, such as level of enjoyment of activities. Objective measures would be more appropriate, for example, in measuring the economic impact of the patient's disability, while subjective measures might be more useful in evaluating a patient's assessment of improvement with treatment.

In selecting a function assessment tool, there is a trade-off between brevity of the instrument and the amount of detail the instrument captures. Instruments range from very brief, global measures, such as the GAF (single scale)<sup>24</sup> and the SDS (3 items),<sup>22</sup> to detailed assessments of multiple functional domains, such as the SF-36 (36 items, 8 domains) and the SAS (48 items, 6 domains).<sup>13</sup> The relative importance of sensitivity versus specificity to the research objective is an important consideration in choosing an instrument for measuring functional impairment. Separate functional domains can be impaired differentially relative to depressive symptoms, and investigators might select an instrument that can show which domains are impaired, capture deficits within a particular domain, or give a global measure that represents functioning across domains, depending on their research goals.

Other issues in measuring function affect study design, selection of the assessment tool, and interpretation of the results. In the interpretation of functional outcomes, attribution of the source of a functional deficit may be complicated by the incidence of comorbidities in the patient population. Functional impairment observed in an MDD clinical trial

can be more directly attributed to the patients' depression if a homogeneous study population was selected by excluding common comorbidities, such as anxiety disorders. The cost of increasing ease of interpretation, however, is a decreasing ability to generalize those findings to other patient populations.<sup>47</sup> Sensitivity to change or to differences between groups is also an important factor in the selection of an assessment tool. Tools such as the SDS, which has demonstrated sensitivity to differences between study drug and placebo, and between study drug and active controls in clinical trials, are essential for treatment outcome studies.<sup>22,33</sup> Finally, measures that will be useful in clinical trials may differ from those that would be useful in the primary care setting. In clinical trials, the instrument is used to help identify drug effects in a homogeneous study population, and generally the mean effect or percent of patients attaining a certain level of response is of interest. In the primary care setting, the assessment should be valid for individual patients, who differ in their specific impairment and personal expectations for treatment, and should be sensitive to change over the course of their therapy.

## Conclusions

More frequently than in the past, patients, clinicians, family members, and third-party payers want evidence that a particular type of treatment is likely to result in improvement of functioning in important life domains. Such evidence is of particular importance when selecting from among a number of different types of treatment demonstrated to be effective in bringing about improvement in symptoms.

The selection of appropriate and practical outcome measures for clinical trials requires careful consideration of the types of functioning expected and the types of patients being studied; that is, one should assess the degree to which specific dimensions of functioning included in a potential assessment procedure are relevant and likely to improve with the treatment. For example, the evaluation of impairment in child care, work, household duties, or other special domains may not be particularly informative in many patients. Depending on the purpose of the planned study, one might want to include patients with specific characteristics, such as those responsible for care of a child under the age of 10 years or expected to function in a particular type of job. If the specific domains included are relevant for a particular type of patient, there should be evidence of adequate reliability and validity, evidence of good discrimination among such patients, and evidence of sensitivity to meaningful improvement in level of impairment in functioning.

There are no set critical domains of functioning that must be evaluated and no optimal number of domains or items per

domain to guide in the selection of a particular assessment tool, although one must take into account the patient burden of all of the assessment procedures being used in the study. Ideally, the procedures selected will be easy for the patients to understand and to complete, the results of the evaluations at baseline and after treatment will be easily communicated, the significance of changes in scores will be easily interpreted, and the evidence provided will be relevant for the choice of future treatment for patients similar to those who participated in the study.

**Author Affiliations:** From the Department of Psychiatry, Columbia University, New York, NY (JE); Advogent, Wayne, NJ (KMD).

**Funding Source:** This study was sponsored by Wyeth, which was acquired by Pfizer Inc in October 2009. Editorial support for this manuscript was provided by Jennifer Karpinski, BA, of Advogent, and was funded by Wyeth.

**Author Disclosure:** Dr Endicott reports the following relationships: Consultant/Advisory Board Member: Cyberonics, Lilly, Wyeth Research; Honoraria: Lilly, Wyeth Research. Dr Dorries is an employee of Advogent, which received funding for the supplement from Pfizer Inc, formerly Wyeth Research.

**Authorship Information:** Concept and design (JE); drafting of the manuscript (JE, KMD); critical revision of the manuscript for important intellectual content (JE, KMD); administrative, technical, or logistic support (KMD); review of literature (JE).

**Address correspondence to:** Jean Endicott, PhD, Professor of Clinical Psychology, Department of Psychiatry, Columbia University, NYSPI-Unit 123, 1051 Riverside Dr, New York, NY 10032. E-mail: je10@columbia.edu.

## REFERENCES

- American Psychiatric Association.** *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision.* Washington, DC: American Psychiatric Association; 2000.
- Hamilton M.** A rating scale for depression. *J Neurol Neurosurg Psychiatry.* 1960;23:56-62.
- Montgomery SA, Åsberg M.** A new depression scale designed to be sensitive to change. *Br J Psychiatry.* 1979;134:382-389.
- Guy W, ed.** Clinical Global Impressions. Publication ADM 76-338. *ECDEU Assessment Manual for Psychopharmacology.* Rockville, MD: US Dept of Health, Education, and Welfare; 1976: 217-222.
- Hirschfeld RM, Montgomery SA, Keller MB, et al.** Social functioning in depression: a review. *J Clin Psychiatry.* 2000;61(4):268-275.
- Keller MB.** Past, present, and future directions for defining optimal treatment outcome in depression: remission and beyond. *JAMA.* 2003;289(23):3152-3160.
- Bech P.** Social functioning: should it become an endpoint in trials of antidepressants? *CNS Drugs.* 2005;19(4):313-324.
- Gelenberg AJ, Thase ME, Meyer RE, et al.** The history and current state of antidepressant clinical trial design: a call to action for proof-of-concept studies. *J Clin Psychiatry.* 2008;69(10): 1513-1528.
- McKnight PE, Kashdan TB.** The importance of functional impairment to mental health outcomes: a case for reassessing our goals in depression treatment research. *Clin Psychol Rev.* 2009;23(3):243-259.
- Frank E, Prien RF, Jarrett RB, et al.** Conceptualization and rationale for consensus definitions of terms in major depressive disorder: remission, recovery, relapse, and recurrence. *Arch Gen Psychiatry.* 1991;48(9):851-855.
- Ballenger JC.** Clinical guidelines for establishing remission in patients with depression and anxiety. *J Clin Psychiatry.* 1999;60(suppl 22):29-34.
- Hirschfeld RM, Dunner DL, Keitner G, et al.** Does psychosocial functioning improve independent of depressive symptoms? A comparison of nefazodone, psychotherapy, and their combination. *Biol Psychiatry.* 2002;51(2):123-133.
- Weissman MM, Bothwell S.** Assessment of social adjustment by patient self-report. *Arch Gen Psychiatry.* 1976;33(9):1111-1115.
- Mulder RT, Joyce PR, Frampton C.** Relationships among measures of treatment outcome in depressed patients. *J Affect Disord.* 2003;76(1-3):127-135.
- Judd LL, Akiskal HS, Zeller PJ, et al.** Psychosocial disability during the long-term course of unipolar major depressive disorder. *Arch Gen Psychiatry.* 2000;57(4):375-380.
- Gallo JJ, Rabins PV, Lyketsos CG, Tien AY, Anthony JC.** Depression without sadness: functional outcomes of nondysphoric depression in later life. *J Am Geriatr Soc.* 1997;45(5): 570-578.
- Lebowitz BD, Pearson JL, Schneider LS, et al.** Diagnosis and treatment of depression in late life. Consensus statement update. *JAMA.* 1997;278(14):1186-1190.
- Solomon DA, Leon AC, Coryell W, et al.** Predicting recovery from episodes of major depression. *J Affect Disord.* 2008;107(1-3):285-291.
- Solomon DA, Leon AC, Endicott J, et al.** Psychosocial impairment and recurrence of major depression. *Compr Psychiatry.* 2004;45(6):423-430.
- Zimmerman M, Ruggero CJ, Chelminski I, et al.** Developing brief scales for use in clinical practice: the reliability and validity of single-item self-report measures of depression symptom severity, psychosocial impairment due to depression, and quality of life. *J Clin Psychiatry.* 2006;67(10):1536-1541.
- Zimmerman M, McGlinchey JB, Posternak MA, Friedman M, Attiullah N, Borescu D.** How should remission from depression be defined? The depressed patient's perspective. *Am J Psychiatry.* 2006;163(1):148-150.
- Sheehan DV, Harnett-Sheehan K, Raj BA.** The measurement of disability. *Int Clin Psychopharmacol.* 1996;11(suppl 3):89-95.
- Williams JBW.** Mental health status, functioning, and disabilities. In: Rush AJ, First MB, Blacker D, eds. *Handbook of Psychiatric Measures.* 2nd ed. Washington, DC: American Psychiatric Publishing, Inc; 2008:83-105.
- Endicott J, Spitzer RL, Fleiss JL, Cohen J.** The global assessment scale. A procedure for measuring overall severity of psychiatric disturbance. *Arch Gen Psychiatry.* 1976;33(6):766-771.
- Ware JE Jr, Sherbourne CD.** The MOS 36-Item Short-Form Health Survey (SF-36). I. Conceptual framework and item selection. *Med Care.* 1992;30(6):473-483.
- Rehm J, Ustun TB, Saxena S, et al.** On the development and psychometric testing of the WHO screening instrument to assess disablement in the general population. *Int J Methods Psychiatr Res.* 1999;8(2):110-122.
- Endicott J, Nee J.** Endicott Work Productivity Scale (EWPS): a new measure to assess treatment effects. *Psychopharmacol Bull.* 1997;33(1):13-16.
- Leon AC, Solomon DA, Mueller TI, Turvey CL, Endicott J, Keller MB.** The Range of Impaired Functioning Tool (LIFE-RIFT): a brief measure of functional impairment. *Psychol Med.* 1999;29(4):869-878.
- Altshuler L, Mintz J, Leight K.** The Life Functioning Questionnaire (LFQ): a brief, gender-neutral scale assessing functional outcome. *Psychiatry Res.* 2002;112(2):161-182.
- Wing JK, Beevor AS, Curtis RH, Park SB, Hadden S, Burns A.** Health of the Nation Outcome Scales (HoNOS). Research and development. *Br J Psychiatry.* 1998;172:11-18.
- Goldstein MJ.** Further data concerning the relation between premorbid adjustment and paranoid symptomatology. *Schizophr Bull.* 1978;4(2):236-243.

32. Rosa AR, Sanchez-Moreno J, Martinez-Aran A, et al. Validity and reliability of the Functioning Assessment Short Test (FAST) in bipolar disorder. *Clin Pract Epidemiol Ment Health*. 2007;3:5-12.
33. 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.
34. Zweig RA, Turkel E. The social adjustment scale-self-report: psychometric properties for older adults. *Psychol Rep*. 2007;101(3 pt 1):920-926.
35. Ramirez A, Ekselius L, Ramklint M. Axis V—Global Assessment of Functioning scale (GAF), further evaluation of the self-report version. *Eur Psychiatry*. 2008;23(8):575-579.
36. Edelman D, Williams GR, Rothman M, Samsa GP. A comparison of three health status measures in primary care outpatients. *J Gen Intern Med*. 1999;14(12):759-762.
37. Page AC, Hooke GR, Rutherford EM. Measuring mental health outcomes in a private psychiatric clinic: Health of the Nation Outcome Scales and Medical Outcomes Short Form SF-36. *Aust N Z J Psychiatry*. 2001;35(3):377-381.
38. Ware J, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care*. 1996;34(3):220-233.
39. Endicott J, Nee J, Harrison W, Blumenthal R. Quality of Life Enjoyment and Satisfaction Questionnaire: a new measure. *Psychopharmacol Bull*. 1993;29(2):321-326.
40. Wisniewski SR, Rush AJ, Bryan C, et al. Comparison of quality of life measures in a depressed population. *J Nerv Ment Dis*. 2007;195(3):219-225.
41. Posl M, Cieza A, Stucki G. Psychometric properties of the WHODASII in rehabilitation patients. *Qual Life Res*. 2007;16(9):1521-1531.
42. Weinstock LM, Miller IW. Functional impairment as a predictor of short-term symptom course in bipolar I disorder. *Bipolar Disord*. 2008;10(3):437-442.
43. Weissman MM, Olfson M, Gameroff MJ, Feder A, Fuentes M. A comparison of three scales for assessing social functioning in primary care. *Am J Psychiatry*. 2001;158(3):460-466.
44. Burns A, Beevor A, Lelliott P, et al. Health of the Nation Outcome Scales for elderly people (HoNOS 65+). *Br J Psychiatry*. 1999;174:424-427.
45. Gowers SG, Harrington RC, Whitton A, et al. Brief scale for measuring the outcomes of emotional and behavioural disorders in children. Health of the Nation Outcome Scales for children and Adolescents (HoNOSCA). *Br J Psychiatry*. 1999;174:413-416.
46. Endicott J, Nee J, Yang R, Wohlberg C. Pediatric Quality of Life Enjoyment and Satisfaction Questionnaire (PQ-LES-Q): reliability and validity. *J Am Acad Child Adolesc Psychiatry*. 2006;45(4):401-407.
47. Klein DF, Thase ME, Endicott J, et al. Improving clinical trials: American Society of Clinical Psychopharmacology recommendations. *Arch Gen Psychiatry*. 2002;59(3):272-278.