

# Assessing Functional Outcomes in Clinical Practice

Michael S. Klinkman, MD, MS

## Introduction

Depression is one of the most common chronic health problems in the United States, with lifetime prevalence estimates of more than 20% in the community.<sup>1,2</sup> Primary care providers manage a substantial proportion of these patients. Prevalence estimates for major depressive disorder (MDD) in primary care practices range from 6% to 14%. “Depression” is listed as a diagnosis in almost 10% of office visits in primary care, and primary care physicians prescribe about half of all antidepressant medications.<sup>3-6</sup> Evidence-based treatment guidelines for the management of depression have been widely disseminated, and current guidelines often incorporate innovations such as the use of measurement-based care and the principle of treatment to remission.

Unfortunately, much of the evidence available is limited regarding assessing the clinical efficacy of interventions during the acute phase of treatment for new episodes of depression. For example, most clinical trials of antidepressant medications evaluate the relative efficacy of active medication versus placebo (or another medication) in reducing depressive symptoms as measured by a standard assessment instrument over a period of between 8 and 16 weeks at the beginning of a treatment episode. To qualify for inclusion in these trials, individuals must meet full diagnostic criteria for MDD and score above a threshold level of severity. Persons with medical and/or mental health comorbidities are, by design, excluded. These design features and inclusion criteria result in a cohort of subjects that bear little resemblance to the majority of depressed patients seen in the primary care setting, who often have chronic depressive symptoms, meet criteria for more than 1 mental health disorder, and have 1 or more chronic medical problems.<sup>7,8</sup> Most of these patients do not fit into the acute-phase treatment framework implicit in clinical guidelines,<sup>9</sup> making it necessary to look beyond short-term efficacy to assess outcomes. Even for those patients who are in acute-phase treatment, the presence of comorbidity makes it difficult to use efficacy and/or remission as the outcome of choice. The following real-world examples from clinical practice can illustrate this point.

- **Case 1.** MJ is a 43-year-old woman with a “simple” depressive episode and no significant comorbidity. Over several weeks of treatment, her 9-item Patient Health Questionnaire (PHQ-9)<sup>10</sup> score has improved from 19 (indicating severe depression) to

## Abstract

Depression is one of the most common chronic health problems in the United States, and primary care providers manage a substantial proportion of these patients. Unfortunately, most current knowledge about treatment effectiveness is limited to the acute phase of treatment for new depressive episodes, although most patients seen in the primary care setting have chronic depressive symptoms, meet criteria for more than one mental health disorder, and have one or more chronic medical conditions. This article examines the shortcomings of our current approach to assessing treatment effectiveness in primary care, despite the availability of good measures of symptom-based recovery, such as the 9-item Patient Health Questionnaire (PHQ-9). Specific emphasis is placed on the need to expand our focus from current symptom-based outcomes of remission and response to measures that can also capture positive emotional recovery, well-being, and functional status, and to integrate these measures into everyday primary care practice. Although there is not yet a standard measure to assess emotional recovery, well-being, or functional recovery, brief measures such as the Sheehan Disability Scale, Quality of Life Enjoyment and Satisfaction Questionnaire, World Health Organization 5-Item Well-Being Index, and new Remission Evaluation and Mood Inventory Tool are available. The opportunity now exists to use these simple tools to integrate outcome monitoring into routine care in the same way other chronic health problems, such as asthma or diabetes, are monitored. Options such as point-of-care outcome assessment with PHQ-9, plus a functional recovery tool; clinician extender (“care manager”) monitoring of depressed patients; or a hybrid approach combining both approaches can be practical and effective.

(*Am J Manag Care.* 2009;15:S335-S342)

For author information and disclosures, see end of text.

■ **Table 1.** Patterns of Comorbidity in the National Comorbidity Study, 2001-2003

<b>12-Month prevalence rates:</b>	
1 threshold mental health disorder only	14.4%
2 threshold disorders	5.8%
3 or more threshold disorders	6.0%
<b>Correlation between presence of major depressive disorder and dysthymia</b>	0.88
<b>Latent class analysis exploring patterns of comorbidity</b>	<ul style="list-style-type: none"> <li>• 3 of 6 identified classes characterized as highly comorbid</li> <li>• About 25% of those with any threshold disorder were highly comorbid</li> </ul>
<p>Respondents receiving care only in the general medical sector had a higher number of chronic physical conditions, less severe mental disorders.  <i>Sources:</i> Kessler RC, et al. <i>Arch Gen Psychiatry.</i> 2005;62(6):617-627.                  Uebelacker LA, et al. <i>Gen Hosp Psychiatry.</i> 2006;28(5):387-395.</p>	

7 (slightly higher than the 5 or less score indicating remission). However, she reports that she does not feel “better” when asked about mood or functional improvement. After increasing her dose of antidepressant medication and referring for short-term talk therapy, her PHQ-9 improves to 4. She is in remission, but still does not feel like she is “better” and has no further functional improvement. By definition, she is in remission, so what are the next steps?

- **Case 2.** JF is a 39-year-old man with a “simple” depressive episode (his first) and no significant comorbidity. After a few weeks of acute-phase treatment, his PHQ-9 score improves from 17 to 12. He says “I’m better, doing great! I’m feeling like this is good enough.” He is not in remission. He does not want to increase medication dosing. What are the next steps when he asks to discontinue medication several weeks later with a PHQ-9 score still at 12?
- **Case 3.** BT is a 48-year-old woman with chronic depression, poorly controlled diabetes and hypertension, chronic pelvic pain, recurrent atypical chest pain (likely anxiety), migraine, and recent cerebrovascular disease leaving her with right leg weakness. She has an unstable home situation, with multiple family members with mood disorders or bipolar disorder. Her antidepressant medication dosing is very difficult due to side effects and poor adherence. Her minimum PHQ-9 score in recent months has been 13. How should treatment effectiveness be assessed? How should “success” be measured if remission is beyond reach, and depression severity scores are high from her chronic medical problems?
- **Case 4.** NC is a 36-year-old woman with minor depression who has met full criteria for MDD in the past. She also meets criteria for panic and generalized anxiety disorder, and carries diagnoses of fibromyalgia, migraine, irritable bowel syndrome, and asthma. Multiple antidepressant medication trials have been prescribed to treat fibromyalgia and mood symptoms with limited impact. Her recent PHQ-9 scores range from 14 to 8, her lowest score ever is 6, but her scores do not correlate with her symptom level or functional status. Her ability to function at work is approaching the crisis point, with

increased absences and decreased productivity, complicated by poor sleep. What is a reasonable measure for response to treatment? What should be considered a “good” clinical outcome for her?

In each of these patients, management of depression is a complicated process, intertwined with treatment of comorbid problems and constrained by patients’ expectations. Our measures for successful or effective care must go beyond the current standards of “recovery” and “remission,” which simply reflect symptom counts of core diagnostic criteria.

This manuscript will explore an expanded approach to assessing recovery that can be applied to all patients seen in everyday primary care practice. Comorbidity and chronicity in primary care patients will be explored, as well as the elements necessary for a clinically meaningful definition of recovery. A short case example of instrument validation will be described, and ways to incorporate functional outcome assessment into routine clinical practice will be discussed.

**Comorbidity in Community and Primary Care Settings**

The National Comorbidity Survey Replication, conducted between 2001 and 2003, compiled 12-month prevalence data for mental disorders from a comprehensive survey of community residents in the United States.<sup>11,12</sup> The key findings of this study are presented in **Table 1**.

Recent data from our ongoing work at the University of Michigan Depression Center confirms the high rate of comorbidity in primary care patients referred for disease management support. Systematic 2-stage screening of all depressed enrollees in the Michigan Depression Outreach and Collaborative Care (M-DOCC) program identified a high proportion meeting criteria for other disorders (**Table 2**): 28% for bipolar disorder, 26% for posttraumatic stress disorder, and 64% for generalized anxiety disorder. A substantial proportion of patients (33%) met criteria for 3 or

■ **Table 2.** Rates of Positive Screens for Comorbid Mental Health Conditions in Depressed Primary Care Patients Referred to Michigan Depression Outreach and Collaborative Care Program, University of Michigan Depression Center, 2006-2007 (n = 210)

Condition	Positive on Primary Screen, %	Positive at Diagnostic Threshold, %
Bipolar disorder	40	28
Generalized anxiety disorder	70	64
Panic disorder	48	41
Posttraumatic stress disorder	33	26
Substance abuse	28	28
Sleep problems	71	N/A

Note: 33% Screened positive for at least 3 comorbid conditions.

Source: Klinkman, unpublished data, University of Michigan Depression Center.

more threshold disorders (unpublished data, courtesy of the University of Michigan Depression Center). This raises the question: What is the primary diagnosis, and does it matter?

### Chronicity in Primary Care Settings

Although this issue receives very little attention in the literature, the large majority of depressed patients seen in primary care have chronic and/or recurring symptoms, often in the context of receiving active treatment. In one recently completed study of active treatment for depression in 109 primary care practices, within a 12-month period, only 1 in 12 treated patients (8%) were entering a new treatment episode: more than 91% were chronic treatment episodes (personal communication, James Gill, MD).

Depression is a chronic disease of varying levels of severity rather than an acute condition—more similar to asthma than appendicitis.<sup>13</sup> As with asthma, depression can be constant and more severe in some cases (similar to chronic asthma), and follow a waxing and waning course in others (similar to intermittent asthma). For those with constant and severe symptoms, detection may be easier, the need for treatment more clear, and response to treatment may be easier to assess. For those in the latter group, it is more difficult to establish a diagnosis and it may be difficult to determine when and for how long to actively treat, as the course of symptoms without treatment may not be easily predictable. As with asthma, active treatment can improve outcomes in both groups. Initial severity of depressive symptoms and initial response to active treatment can be used to guide treatment decisions and establish long-term prognosis.

### Problems With “Remission” and “Recovery” as Primary Outcomes in Primary Care

The current goal of active treatment is to achieve remission, defined as the absence of depressive symptoms and

operationally defined in primary care clinical trials as a PHQ-9 score of 5 or less. There are at least 3 problems with the use of remission rate as the primary outcome measure for effective treatment.

First, as seen in our clinical examples, remission often does not track along with the patient’s own sense of clinical improvement. Remission rates in state-of-the-art clinical trials are low, and many patients who feel as though they have returned to their “normal self” fall short of remission criteria. For example, in the Sequenced Treatment Alternatives to Relieve Depression study, the cumulative remission rate after 4 stages of treatment was 67% for a cohort of patients with limited comorbidity entering a new treatment episode.<sup>14</sup> Other primary care interventions report lower remission rates,<sup>15</sup> and longer-term studies show sustainable remission rates over time only if patients continue to receive active management at high levels of intensity.<sup>16,17</sup> From the clinician’s point of view, a standardized outcome measure that contradicts the patient’s own report will create dissonance in the clinician’s mind, and will raise questions about the real-world validity of the measure.

Second, achievement of remission is confounded by the presence of significant comorbidity. The somatic symptoms included as criteria for remission (fatigue, sleep problems) are common, and severe, in other chronic mental health and medical conditions. In many cases, these symptoms are not from depression, cannot be reduced by depression treatment, and anchor the scores on assessment tools so that a “remission” score cannot be obtained. Almost no evidence on achievable remission rates under these real-world conditions exists. In our most recent trial of primary care depression management, we achieved a remission rate of 49.2% at 18 months (compared with 27.2% for usual care) in a cohort of chronically depressed, highly comorbid patients,<sup>18</sup> suggesting that, at best, half of these patients might reach “remission” as currently defined.

■ **Figure 1.** 2 × 2 Array of Criterion-Based Depression Symptoms (“Depression Symptoms”) Versus Other Emotional/Functional Aspects of Recovery (“Other Symptoms”)

		Depression Symptoms	
		Above threshold (not in remission)	Below threshold (in remission)
Other Symptoms	Above threshold	<b>A</b>	<b>B</b>
	Below threshold	<b>C</b>	<b>D</b>

Finally, the concept of remission itself is insufficient to define outcomes for a chronic condition. We do not speak of “remission” from asthma or diabetes, even when clinical indicators such as peak expiratory flow rate or hemoglobin A1C normalize. For these and other chronic medical conditions, we also look at functional status or health-related quality-of-life (HRQOL) measures<sup>19-21</sup> to provide a more complete assessment of the effectiveness of care.

The most common secondary outcome measure in depression clinical trials is recovery, operationally defined as a 50% or greater reduction in severity as measured by a standard measure such as the PHQ-9, 17-item Hamilton Rating Scale for Depression,<sup>22</sup> or the Quick Inventory of Depression Symptomatology–Self-Report.<sup>23</sup> This measure is also difficult to apply in primary care. For patients who “re-enter” active treatment with partially treated symptoms, an initial PHQ-9 score may be sufficiently low (eg, 9) to make 50% improvement a more stringent outcome measure than remission. More importantly, this measure is based on the same measurement tool as remission and cannot assess the important domains of functional status or other HRQOL measures.

**Expanded Definition of Recovery for Primary Care—Remission Plus Function**

Figure 1 shows how we might figuratively place patients being treated for depression in a 2 × 2 array based on their remission status (in columns) and their improvement in other unmeasured aspects of improvement such as functional status (in rows). Patients in the left column (groups A and C), with standard depression measure scores above a threshold, are not in remission, whereas those in the right column (B and D) with scores below threshold are in remission.

For groups A and D, criterion-based remission scores are congruent with other aspects of improvement and a remission measure can correctly categorize clinical outcome. However, things are more complicated for those in the “off-diagonal” groups B and C. Patients in group B, in remission, may still experience considerable dysfunction; recall patient MJ, in case 1. Patients in group C, not in remission, may be highly functional and may consider themselves to be recovered; recall patient JF in case 2. Think also of patients BT and NC (cases 3 and 4), whose chronicity and comorbidity makes it highly unlikely that they will ever be in group B or D; their best possible outcome is to be included in group C.

The 4 groups are important in developing a more complete picture of relevant primary care outcomes of depression treatment. It is important to be able to discriminate between remission in symptom-based criteria and full recovery (B vs D), and to know when chronically depressed patients achieve a high level of functional recovery (C vs A).

This expanded approach to measuring depression outcomes is consistent with current expert opinion,<sup>24</sup> and patients seem to intuitively understand the importance of this expanded approach. Zimmerman and colleagues<sup>25</sup> surveyed 535 psychiatric outpatients to identify items they believed were most important in determining whether their depression was in remission; the highest-rated items were related to positive mental health (optimism, self-confidence), return to normal level of functioning, and feeling like one’s usual normal self.

These items can be loosely grouped into “feeling” and “doing” categories. *Feeling* refers to a personal sense of positive emotional recovery, return of resilience, or improvement in one’s sense of well-being: “Am I feeling better (good) about myself? Am I less fragile? Is my emotional status stable?” *Doing* refers to improvement in what one is able to do,

■ **Figure 2.** 2 × 2 Array: Proportion of Patients Within Cells of Criterion-Based Symptoms (PHQ-9) Versus Other Emotional/Functional Aspects of Recovery (REMIT) in Cross-Sectional Survey of Primary Care Patients (n = 1006)

		Criterion-Based Depression Score (PHQ-9)	
		PHQ-9 >5 (not in remission)	PHQ-9 ≤5 (in remission)
REMIT Score	Upper half ("Worse")	63.7%	5.7%
	Lower half ("Better")	36.3%	94.3%

PHQ-9 indicates 9-item Patient Health Questionnaire; REMIT, Remission Evaluation and Mood Inventory Tool.  
Source: Nease et al, unpublished data from REMIT validation study, 2008.

in both cognitive and physical performance areas, rather than whether one is satisfied about it: "Is my work performance improving? Am I keeping up with daily household tasks? Am I able to carry out social interactions?"

Developing measures that effectively capture these additional categories has been challenging. The items are difficult to measure as they are subjective by definition, related to an individual's internal sense of what is "normal" for them. They also are subject to ceiling effects in measurement, such as the presence of comorbid medical illness that limits potential gains in functional status. A large inventory of patient self-report measures of overall functional status, well-being, and HRQOL already exists (a good example is the Short Form-12), but these instruments are quite unwieldy for use in everyday primary care practice.<sup>26</sup> Brief measures of well-being and functional status such as the Sheehan Disability Scale (SDS),<sup>27</sup> Quality of Life Enjoyment and Satisfaction Questionnaire (Q-LES-Q),<sup>28</sup> and World Health Organization 5-Item Well-Being Index (WHO-5)<sup>29</sup> have seen some use in clinical trials,<sup>30-33</sup> but their lack of specificity limits their usefulness in highly comorbid primary care patients. Our challenge is to identify the best items from these instruments and blend them with a few new items to create a brief, simple tool to pair with symptom-based measures such as the PHQ-9.

#### Case Study: Development of the REMIT Tool— A Measure of Nonsymptom Recovery

Our recent experience in developing the Remission Evaluation and Mood Inventory Tool (REMIT) can serve to illustrate the challenges of instrument development.<sup>34</sup> Our

primary objective was to develop a brief, clinically useful tool to measure the most important nonsymptom aspects of recovery as a supplement to a standard measure such as the PHQ-9. We carried out a literature review and conducted secondary analyses of existing longitudinal studies of primary care patients to identify the most promising domains and individual items that were conceptually and empirically: (1) related to patients' own sense of recovery, (2) not highly correlated with symptom-based measures, and (3) followed a different trajectory of improvement over time than symptom-based measures. A preliminary list of 26 items covering 8 domains was revised and expanded through a series of consultations with experts in the field. A final list of 16 candidate items was tested in a cross-sectional survey of 1000 depressed primary care patients in Michigan and Indiana.

Our analyses identified a core set of 7 items weakly correlated with PHQ-9 score and strongly correlated with the patient's own self-assessment of recovery. The PHQ-9 score explained about 50% of the variance in self-assessed recovery, whereas the 7 REMIT items explained an additional 10% of the variance. We found that these items were strongly related to recovery, but in a different pattern, in patients with significant pain. In constructing a version of the 2 × 2 array previously described, we found that more than one third (36.3%) of depressed patients who had not achieved remission scored in the lower (better) half on the 7 items (Figure 2). This corresponds to group C, a clinically important group that includes patients JF, BT, and NC.

The current REMIT (Figure 3) consists of 7 items covering the domains of positive recovery, emotional control, functional status, and resilience, plus 2 "pain" items. We are

■ **Figure 3.** Current Version of Remission Evaluation and Mood Inventory Tool

For each question, please **CIRCLE THE NUMBER** that corresponds to your answer.  
Over the *last 2 weeks* . . .

	None of the time	A little of the time	Some of the time	Most of the time	All of the time	
1. Did you feel happy?	4	3	2	1	0	— +
2. Did you feel content?	4	3	2	1	0	— +
3. Did you feel in control of your emotions?	4	3	2	1	0	— +
4. Did you bounce back when things went wrong?	4	3	2	1	0	— +
5. Did you feel nervous, anxious, or on edge?	0	1	2	3	4	— +
6. Did you isolate yourself from others around you?	0	1	2	3	4	— +
7. Did the future seem dark to you?	0	1	2	3	4	— +
	Not at all	A little bit	Somewhat	Moderately so	Very much so	
8. Were you bothered by aches and pains?	0	1	2	3	4	— +
9. Were you bothered by other physical symptoms besides pain?	0	1	2	3	4	—

**TOTAL =**

currently evaluating the usefulness of this tool as a supplement to the PHQ-9 in routine patient care.

**Putting It All Together: Monitoring Emotional (“Feeling”) and Functional (“Doing”) Outcomes in Primary Care**

Even if we can identify, or create, a practical tool to measure extended outcomes in primary care, we face a second challenge in effectively monitoring outcomes in everyday practice. The guiding principle for monitoring should be “simple tools, used flexibly,” but the development of an organized system of care that utilizes these tools is very important. There are 2 basic ways to implement measurement-based care for depression: point-of-care assessment and clinician extender monitoring and feedback. The simplest approach is to carry out point-of-care assessment by having patients complete a symptom-based tool along with a supplemental tool at the beginning of an office visit, while waiting to see the clinician. This could be administered by a medical assistant or nurse in the same way hemoglobin A1C or blood pressure measurements are obtained prior to the visit, or it could be self-administered prior to the visit. The completed instrument would be reviewed and scored during the encounter, then used to decide on any changes in management. This approach is relatively easy to implement and inexpensive, ensures that

care for depression remains on the agenda at routine office visits even during the maintenance phase of treatment, and fosters focused and efficient discussion of management. Its primary limitation is its linkage to a visit: patients who cancel or do not schedule appointments cannot be monitored, and this is a common problem in depression treatment.

Clinician extender monitoring and feedback is carried out parallel to primary care treatment. It usually involves a series of telephone calls initiated by a care manager, who may be an agent of the practice, provider organization, disease management program, or an insurance company. In most programs, the care manager calls a patient at regular intervals during the acute phase of treatment, provides tangible education or support during each call, monitors severity of symptoms, and feeds information back to the referring primary care clinician. In some settings, the process is automated through outbound calls employing an interactive voice recognition system to capture and forward patient responses. This approach enhances care by providing additional points of contact at a time when patients may be most ambivalent about treatment, and care managers can provide patient education and support beyond the level available in the primary care office. However, these programs also share significant limitations: they are more expensive to design and implement (and very difficult to sustain), their integration into practice workflow

may be poor (so that clinicians cannot use the feedback at the point of care), and they largely operate as disease-specific programs outside the context of patients' other medical care needs. Most importantly, once past the acute phase of treatment, monitoring may end or frequency may be reduced to the point of being ineffective. Most current programs provide support only for the acute phase of treatment, which would not address the needs of patients BT or NC.

A hybrid model that supplements point-of-care assessment with clinician extender monitoring may prove to be the best option. In this approach, some patients with uncomplicated depression or a mild episode could be managed with point-of-care support. Others with more complex or chronic depression could be followed by a care manager who can provide additional support, more frequent monitoring and feedback during acute-phase treatment, and less frequent monitoring during maintenance phase treatment when many patients are lost to follow-up. This approach can fill the gaps inherent in point-of-care assessment while making more efficient use of the care manager.

We have experience with the hybrid approach in the M-DOCC program previously described. Primary care physicians at some sites now have the option of using point-of-care outcome assessment for all depressed patients while continuing to refer patients at their discretion to M-DOCC for care manager monitoring and feedback. Where this option is available, referrals of the least complex patients have declined significantly, whereas referral rates for more complex patients remained constant. We have been able to extend 1 care manager to cover 5 large primary care practice sites, enhancing the efficiency of the program.

## Conclusions

Primary care physicians provide care for a wide range of depressed patients. Relatively few patients are beginning acute-phase treatment for new-onset MDD; most have complex presentations of depression and comorbid health problems or are chronically depressed. We need to expand our focus from the symptom-based outcomes of remission or response to measures that also can capture positive emotional recovery, a sense of well-being, and functional status, and we need to integrate these measures into everyday primary care practice.

Good measures of symptom-based recovery are available for use in primary care, with the PHQ-9 emerging as a standard. We are still working to find or develop an equivalent standard for the expanded "feeling" and "doing" aspects of recovery, but instruments such as the SDS, Q-LES-Q, WHO-5, or the new REMIT may provide a reasonable first pass at a simple and complementary tool.

Developing a system that can integrate outcome monitoring into routine care presents a greater challenge, especially for practices without access to sufficient resources to create or link to a formal clinician extender/care manager program. One practical approach for these practices might be to: (1) carry out point-of-care outcome assessment with the PHQ-9 plus SDS, Q-LES-Q, WHO-5, or REMIT for all depressed patients; and (2) identify a practice staff member (eg, an office nurse) who could create a simple depression disease registry, contact patients in the registry at intervals based on their risk or time since last follow-up, and do telephone monitoring using the same tools. This approach should help both patient and clinician maintain a longer-term focus on those outcomes of most importance to everyday life.

**Author Affiliations:** From the Departments of Family Medicine and Psychiatry, University of Michigan Health System, Ann Arbor.

**Funding Source:** This study was sponsored by Wyeth, which was acquired by Pfizer Inc in October 2009. Medical writing and editorial support for this manuscript was provided by Steven J. Cally, PhD, and Nicole Rodrigues, MS, of Advogent, and was funded by Wyeth.

**Author Disclosure:** The author reports the following relationships: Scientific Advisory Board/Honoraria: Wyeth Pharmaceuticals; Medical Advisory Board: Cielo MedSolutions, LLC.

**Authorship Information:** Concept and design; acquisition of data; analysis and interpretation of data; drafting of the manuscript; and critical revision of the manuscript for important intellectual content.

**Address correspondence to:** Michael S. Klinkman, MD, MS, Associate Professor, Departments of Family Medicine and Psychiatry, University of Michigan Health System, Ann Arbor, MI 48104-1213. E-mail: mklinkma@umich.edu.

## REFERENCES

1. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005;62(6):593-602.
2. Kessler RC, McGonagle KA, Zhao S, et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. *Arch Gen Psychiatry*. 1994;51(1):8-19.
3. Barrett JE, Barrett JA, Oxman TE, Gerber PD. The prevalence of psychiatric disorders in a primary care practice. *Arch Gen Psychiatry*. 1988;45(12):1100-1106.
4. Katon W. The epidemiology of depression in medical care. *Int J Psychiatry Med*. 1987;17(1):93-112.
5. Leon AC, Olfson M, Broadhead WE, et al. Prevalence of mental disorders in primary care. Implications for screening. *Arch Fam Med*. 1995;4(10):857-861.
6. Mojtabai R. Residual symptoms and impairment in major depression in the community. *Am J Psychiatry*. 2001;158(10):1645-1651.
7. Zimmerman M, Mattia JI, Posternak MA. Are subjects in pharmacological treatment trials of depression representative of patients in routine clinical practice? *Am J Psychiatry*. 2002;159(3):469-473.
8. Zimmerman M, Chelminski I, Posternak MA. Generalizability of antidepressant efficacy trials: differences between depressed

psychiatric outpatients who would or would not qualify for an efficacy trial. *Am J Psychiatry*. 2005;162(7):1370-1372.

9. **Katz MM.** Clinical trials of antidepressants: time to shift to a new model. *J Clin Psychopharmacol*. 2008;28(4):468-470.

10. **Kroenke K, Spitzer RL, Williams JB.** The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16(9):606-613.

11. **Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE.** Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005;62(6):617-627.

12. **Uebelacker LA, Wang PS, Berglund P, Kessler RC.** Clinical differences among patients treated for mental health problems in general medical and specialty mental health settings in the National Comorbidity Survey Replication. *Gen Hosp Psychiatry*. 2006;28(5):387-395.

13. **Klinkman MS, Schwenk TL, Coyne JC.** Depression in primary care—more like asthma than appendicitis: the Michigan Depression Project. *Can J Psychiatry*. 1997;42(9):966-973.

14. **Rush AJ, Trivedi MH, Wisniewski SR, et al.** Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR\*D report. *Am J Psychiatry*. 2006;163(11):1905-1917.

15. **Kroenke K, West SL, Swindle R, et al.** Similar effectiveness of paroxetine, fluoxetine, and sertraline in primary care: a randomized trial. *JAMA*. 2001;286(23):2947-2955.

16. **Wells K, Sherbourne C, Schoenbaum M, et al.** Five-year impact of quality improvement for depression: results of a group-level randomized controlled trial. *Arch Gen Psychiatry*. 2004;61(4):378-386.

17. **Katon W, Russo J, Von Korff M, et al.** Long-term effects of a collaborative care intervention in persistently depressed primary care patients. *J Gen Intern Med*. 2002;17(10):741-748.

18. **Klinkman M, Grazier KL, Emptage N, et al.** First results from the depression in primary care demonstration project: integrating “disease” management into the primary care process. *Fam Med*. 2006;38(suppl 1).

19. **Juniper EF, Buist AS, Cox FM, Ferrie PJ, King DR.** Validation of a standardized version of the Asthma Quality of Life Questionnaire. *Chest*. 1999;115(5):1265-1270.

20. **Bradley C, Todd C, Gorton T, Symonds E, Martin A, Plowright R.** The development of an individualized questionnaire measure of perceived impact of diabetes on quality of life: the ADDQoL. *Qual Life Res*. 1999;8(1-2):79-91.

21. **Garin O, Ferrer M, Pont A, et al.** Disease-specific health-related quality of life questionnaires for heart failure: a systematic review with meta-analyses. *Qual Life Res*. 2009;18(1):71-85.

22. **Hamilton M.** A rating scale for depression. *J Neurol Neurosurg Psychiatry*. 1960;23:56-62.

23. **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.

24. **Keller MB.** Past, present, and future directions for defining optimal treatment outcome in depression: remission and beyond. *JAMA*. 2003;289(23):3152-3160.

25. **Zimmerman M, McGlinchey JB, Posternak MA, Friedman M, Attiullah N, Boerescu D.** How should remission from depression be defined? The depressed patient’s perspective. *Am J Psychiatry*. 2006;163(1):148-150.

26. **Ware J Jr, 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.

27. **Sheehan DV.** Sheehan Disability Scale. In: **Rush AJ et al, eds.** *Handbook of Psychiatric Measures*. First ed. Washington, DC: American Psychiatric Association; 2000:113-115.

28. **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.

29. **World Health Organization.** *Wellbeing Measures in Primary Health Care/the DepCare Project*. Copenhagen, Denmark: WHO Regional Office for Europe; 1998.

30. **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.

31. **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.

32. **Henkel V, Mergl R, Kohnen R, Allgaier AK, Moller HJ, Hegerl U.** Use of brief depression screening tools in primary care: consideration of heterogeneity in performance in different patient groups. *Gen Hosp Psychiatry*. 2004;26(3):190-198.

33. **Newnham EA, Hooke GR, Page AC.** Monitoring treatment response and outcomes using the World Health Organization’s Wellbeing Index in psychiatric care. *J Affect Disord*. 2009 [epub ahead of print]

34. **Nease DE Jr, Aikens JE, Kroenke K, Klinkman MS.** The remission evaluation and mood inventory tool project - toward a broader measure of depression remission for primary care. *J Royal Australian N Z Coll Psychiatr*. 2007;41:A319.



