Definitions of Antidepressant Treatment Response, Remission, Nonresponse, Partial Response, and Other Relevant Outcomes: A Focus on Treatment-Resistant Depression

Andrew A. Nierenberg, M.D., and Lindsay M. DeCecco, B.A.

Multiple definitions have been used to characterize the outcome of treatments for depression. Beyond the simple criterion of a statistically significant improvement in depression rating scale scores, researchers have had to use more clinically relevant categorical outcomes: response (without remission), remission, nonresponse, partial response, relapse, recurrence, recovery, and, more recently, depressive breakthrough. This article reviews the definitions of these terms and their relevance for the study of treatment-resistant depression. *(J Clin Psychiatry 2001;62[suppl 16]:5–9)*

Since the advent of randomized clinical trials designed for the assessment of efficacy of psychopharmacologic agents, the psychiatric research community has had to measure change in depression. But researchers have used multiple definitions of change to test the advantage (or lack thereof) of an antidepressant over placebo; definitions range from a statistically significant difference in absolute measures on depression scale scores (from baseline to endpoint) to a more clinically based categorical set of definitions of response, nonresponse, partial response, remission, and recovery.[1–3] Remarkably, although the U.S. Food and Drug Administration (FDA) and other regulatory bodies demand that standard scales be used for research, much of the clinical community has been reluctant to adopt numerate assessments of psychopathology, perhaps contributing to the disquieting realization that a gap exists between research findings and clinical practice.[4,5] This gap is particularly evident in the literature on treatment-resistant depression because so many different definitions are used to assess nonresponse.[6,7] This article reviews the widely accepted definitions of categories of response used in clinical research, with a focus on methods used to study treatment-resistant depression.

From the Depression Clinical and Research Program and Bipolar Programs, Massachusetts General Hospital; and Harvard Medical School, Boston.

Presented at the planning teleconference "Recognizing Treatment-Resistant Depression," held November 6, 2000, and supported by an unrestricted educational grant from Eli Lilly and Company.

Reprint requests to: Andrew A. Nierenberg, M.D., WAC 815, Depression Research Program, Clinical Psychopharmacology Unit, Massachusetts General Hospital, 15 Parkman St., Boston, MA 02114 (e-mail: anierenberg@partners.org).

CLINIMETRICS AND THE MEASUREMENT OF DEPRESSION

The measurement of clinical phenomena has been an integral part of medicine since antiquity, but reliable and valid numerate measures are usually reserved for the province of clinical research.[8] It is through measurement that we know whether our treatments work. But psychopathology is particularly challenging to measure, because much of the raw data used to assess change are patient-reported, a primary source that, in turn, can be influenced by the very psychopathology that is being measured, especially for depression and all of its attendant cognitive distortions.

Since the psychopharmacologic revolution of the 1950s, the worldwide metric of choice to measure depression in research settings has been the Hamilton Rating Scale for Depression (HAM-D).[9,10] The HAM-D has gone through many transformations and modifications, including its original 17-item and expanded 21-, 24-, 25-, 28-, and 32-item versions, in addition to a widely used structured interview guide for seasonal affective disorder.[11] As originally designed by Max Hamilton, the scale is supposed to be filled out by the interviewer and an observer after an extensive unstructured clinical interview, with the sum of their results as the final score. If no observer is available, the single score is supposed to be doubled.[9] Needless to say, the scale is rarely used as originally designed. Of particular interest is the fact that the scale was developed for inpatients (usually with severe depression and melancholia), and its applicability to outpatients may be limited. Nonetheless, the HAM-D has remained the gold standard for over 40 years and is currently reported as the main outcome variable in the majority of clinical trials in depression. The HAM-D has been dissected, analyzed, and criticized dur-

Figure 1. Estimated Proportion of Patients Who Meet Response Criteria in Typical Antidepressant Trials

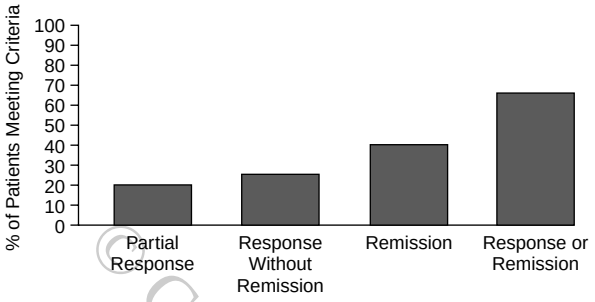
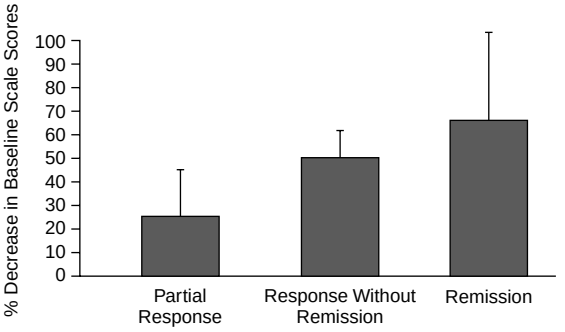


Figure 2. Categories of Response: Relationship to Decrease in Baseline Depression Scale Scores[a]



[a]Vertical line = range of % decrease to qualify as category.

ing its reign[12]; its hegemony is only now being challenged by such scales as the Montgomery-Asberg Depression Rating Scale[13] and the Inventory for Depressive Symptomatology.[14,15] These newer scales, as well as a 6-item subscale of the HAM-D,[16] have been found to be more sensitive to change than the original 17-item HAM-D.

No categorical decisions about response need be made when a clinical trial dictates that efficacy will be determined by a statistically significant decrease in depression rating scale scores after treatment. But researchers do need to decide the parameters that define categories of response for the results of the study to be relevant for clinicians. For research of treatment resistance in particular, it is critical to define nonresponse in order to test treatments for nonresponding patients.

REMISSION

To achieve remission is to be "normal," free of depressive symptoms. What is most striking is that, even using the original 17-item HAM-D, researchers have reported several cutoff scores, from 7 to 10, to define remission. Even with a cutoff point of ≤ 7, residual symptoms still appear to be present.[17] The clinical implication of remission is that only those patients who achieve remission improve fully, not only in terms of symptoms, but also in terms of functioning. Although it is a worthwhile goal of treatment, only 25% to 50% of patients in clinical trials achieve remission[18] (Figure 1). Furthermore, remission is usually assessed at a single timepoint (the end of the trial) and is reported without any requirement that it be sustained for any length of time (see "Recovery").

RESPONSE

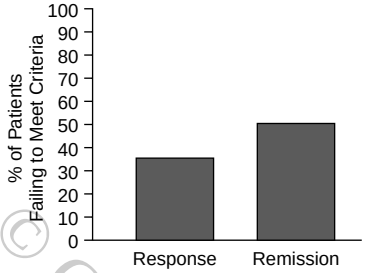
It is widely known within the clinical research community of depression researchers that a response is defined as a ≥ 50% decrease from baseline depression scale scores to trial endpoint[1,3] (Figure 2). The 17-item HAM-D is the most cited scale that requires this level of decrease to define the response category, and it is this version of the HAM-D that omits what are commonly known as reverse neurovegetative signs, i.e., hypersomnia and hyperphagia. This means that if depressed patients present with atypical depression, some of their most troubling symptoms are neither measured at baseline nor included among the symptoms that define response. How a ≥ 50%, as compared with a 40% or 60%, decrease from baseline measures became the standard definition is unclear. The problem with using the criterion of a ≥ 50% drop is that if a patient starts a trial with a baseline HAM-D score in the severe range, 32 for example, then the patient can have a final score that, while defined as a response (16, in this example), would still qualify as a rating that would allow the patient to enter the study. In other words, response without remission, while a clinically important improvement, results in many patients' having substantial residual symptoms.

An alternative set of measures accepted for clinical trials that are more clinically intuitive and sensible are the Clinical Global Impressions-Severity of Illness (CGI-S) and -Improvement (CGI-I) scales.[19] The CGI-S is a simple 7-point scale that describes severity as an overall impression, integrating symptoms and functioning, but without any clear anchor points. The CGI-I categorizes the degree of improvement relative to baseline, ranging from very much improved to very much worse. Using the CGI-I, a rating of "much improved" or "very much improved" is frequently used as a definition of response. The advantage of using this scale is that clinicians can readily understand the CGI-I; the disadvantage is that it lacks precision and anchor points, making it difficult to generalize from the results of one clinician to the results of another.

PARTIAL RESPONSE

Somewhere below response and above nonresponse is the patient who is better than at baseline but clearly continues to be symptomatic, such that further treatment is

Figure 3. Estimated Proportion of Patients Who Fail to Meet Response or Remission Criteria in Typical Antidepressant Trials



required. This category, partial response, has been most frequently defined as < 50% but ≥ 25% decrease from baseline depression scale scores[6] (Figure 2) and corresponds to a CGI-I score of "minimally improved." The patients who fall into this category can either continue with their current treatment (perhaps with an augmentation strategy) or switch to another treatment. Although it would appear simple to stop the treatment that led to a partial response and go on to the alternative treatment, switching risks losing even the slight benefit associated with a partial response—a benefit that can be appreciated by the patient only when it is taken away by stopping treatment.

NONRESPONSE

It is self-evident that the field of treatment-resistant depression rests on the definitions of response and nonresponse.[6] Yet no standard definition has been used across studies.[7] In one sense, it is a simple matter of where one decides to draw the line. In another sense, it is the details that make the decision about how to define nonresponse more complicated than one would expect at first glance. On one end of the spectrum, nonresponse can be defined as failing to achieve a minimal partial response, e.g., < 25% decrease from baseline score as measured by the HAM-D. In the middle is the criterion for failing to achieve a response, e.g., < 50% decrease from baseline. At the far end is the criterion for failing to achieve remission (remission defined as, e.g., final HAM-D score ≤ 7). Each choice carries its own advantages and disadvantages. Defining nonresponse as failing to achieve a minimal decrease in scale score (< 25%) will include only those patients who fail to benefit at all and exclude those who, while improved, continue to have substantial residual symptoms. To include those who fail to respond (response defined as ≥ 50% decrease) leaves out those who respond but fail to remit and fail to reach full symptomatic and functional improvement. To include those who fail to remit raises the bar to such heights that this criterion would include the majority of patients treated within the usual clinical time frames.[18] Using the criterion of failing to achieve remission would, therefore, include over half of patients as meeting criteria for failing treatment (Figure 3).

The trend is transparent: increase the threshold of improvement to define failure of treatment and increase the proportion of treatment failures. On a more clinical level, however, it is a reasonable treatment goal for patients to be completely well and as close to normal as possible.[18,20] It is for this reason that the unprecedented National Institute of Mental Health (NIMH) contract study Sequenced Treatment Alternatives to Relieve Depression (STAR*D)[21] uses failure of remission as the threshold to define treatment failure. The challenge is that many depressed patients have comorbid conditions and/or side effects from medications that can prevent them from achieving a symptom-free state. Increasing evidence shows that antidepressants and specific psychotherapies combined result in higher remission rates than either alone for more severely depressed patients and patients with chronic depression.[22,23]

RESIDUAL SYMPTOMS

Successfully treated depressed patients, both those who respond but fail to remit and those who remit, can continue to have not only residual attenuated depressive symptoms, but also persistent symptoms not usually considered among the core symptoms of depression (e.g., irritability), problems with depressive thinking, and problems with functioning socially and at work.[24–27] Residual attenuated depressive symptoms can include insomnia (early, middle, or terminal), fatigue, psychic and somatic anxiety, excessive reactivity to social stress, pessimism, and mild dysphoria.[17,25] Patients can also take a considerable amount of time after improvement to redevelop their interests in activities and their motivation to follow up on those interests. In the parlance of structured questions for the HAM-D, patients may continue to feel that they have to push themselves to do things. Furthermore, they may also feel that the experience of pleasure, while not completely absent, remains blunted.

Other aspects of depression that can persist after response to treatment are dysfunctional attitudes and depressive cognitions. Patients with a tendency for perfectionism continue to be overly self-critical and fail to cope successfully with external criticism. They can readily experience guilt and attribute blame to themselves, with a tendency to lack self-forgiveness. These persistent qualities can lead to continual problems with interpersonal relationships and with work. It is not surprising that improvements in interpersonal functioning can lag behind symptomatic improvement. If a patient has been depressed for a substantial period of time, then people involved with that patient have to cope with the depressed patient's withdrawal, lack of motivation, negativity, passivity, and irritability. Improvement can be met with skepticism and wariness, and perhaps even with some resentment. It is only after pa-

tients have shown that their improvement has consolidated and persists that their loved ones may feel that it is safe to reengage with them. Yet even with long periods of sustained recovery, depressed patients have been found to have higher scores on measures of dependency than never-depressed controls.[28] It is unknown whether this dependency trait may have predisposed the patients to become depressed in the first place, is associated with conditions comorbid with depression, or is a psychological scar of the depression.

The great importance of residual symptoms of depression is that they put patients at high risk of relapse and recurrence. As part of the large NIMH Collaborative Depression Program, naturalistic follow-up showed that, after recovery from major depression, residual subsyndromal depression was associated with an odds ratio of 3.5 for subjects with subsequent relapse compared with those who had a full acute recovery.[29] This higher risk of relapse associated with residual symptoms is greater than the well-known risk associated with ≥ 3 prior depressive episodes. A similar finding was found in a follow-up study of responders to cognitive therapy.[30]

RECOVERY

Now that the definitions of response without remission, remission, relapse, and recurrence have been described, the next alliterative category is recovery. The most commonly used definition of recovery from a depressive episode is an 8-week period of no longer meeting DSM criteria for major depression.[31] This time period is important for the counting of episodes. Using this definition, it is immediately apparent that patients can have substantial residual symptoms persist while meeting the criterion for recovery. For example, according to the DSM diagnostic schema, patients meet criteria for major depression when they have 5 of 9 definite depressive symptoms. This means that patients can have 4 symptoms, no longer meet criteria, and be considered recovered. An alternative to the criterion of no longer having 5 symptoms is that patients have no or minimal definitive symptoms.[31] As for minimal duration necessary to define recovery, it is unclear how 8 weeks of wellness became the standard for separating depressive episodes. But a minimum of 8 weeks seems eminently reasonable as an arbitrary definition of recovery from a single episode of depression.

SUSTAINED RECOVERY AND DEPRESSIVE BREAKTHROUGH

More problematic is how to define recovery from recurrent depression. How long must one be free of relapses or recurrences before one has sustained recovery from recurrent depression? Perhaps in this case, it is best to turn to other disorders/diseases for analogy: depression, like hypertension, is not cured and one does not recover from it; instead, the manifestations of the disorder can be controlled and managed over the long term. Patients with major depression tend to have multiple episodes; episodes that reappear within 6 months of acute response are called *relapses*, whereas those that occur after 6 months are called *recurrences*.[1,3] Theoretically, relapses are considered a return of the original episode, whereas recurrences represent a new episode; however, these are hypotheses with no substantiating data. The majority of patients who do not take long-term antidepressant treatment experience depressive relapses or recurrences.[32] Studies that use a placebo-substitution paradigm and randomly assign depressed patients who responded to antidepressants to either continue on the antidepressant or switch to placebo reveal that antidepressants lower the risk of relapse and recurrence.[33–38] Nonetheless, observational studies have shown that, even while taking long-term antidepressants for prophylaxis, 20% to 80% of patients develop another depressive episode within 1 to 5 years after an acute response.[32,39,40] A succinct term for relapse or recurrence during long-term antidepressant treatment is *depressive breakthrough*.[41] Of great concern are patients who develop a recurrence and then recover from the second depressive episode; they are at risk of developing even more episodes.[32] Furthermore, depressive relapse and recurrence rates may be substantially higher in clinical, as compared with research, populations.

CONCLUSIONS

The optimal outcome for an individual patient with recurrent depression is to recover from an episode and never become depressed again. So far, while antidepressants and structured psychotherapies reduce the likelihood of depressive breakthrough, no treatment is perfect. The longer one waits, the more likely a patient with recurrent depression will have yet another episode. Future research on treatment-resistant depression will focus not only on acute response and remission, but also on the long-term prevention of relapse, recurrence, and depressive breakthrough.

Disclosure of off-label usage: The authors have determined that, to the best of their knowledge, no investigational information about pharmaceutical agents has been presented in this article that is outside U.S. Food and Drug Administration–approved labeling.

REFERENCES

1. Frank E, Prien RF, Jarrett RB, et al. Conceptualization and rationale for consensus definitions of terms in major depressive disorder. Arch Gen Psychiatry 1991;48:851–855
2. Prien RF, Carpenter LL, Kupfer DJ. The definition and operational criteria for treatment outcome of major depressive disorder. Arch Gen Psychiatry 1991;48:796–800
3. Riso LP, Thase ME, Howland RH, et al. A prospective test of criteria for response, remission, relapse, recovery, and recurrence in depressed patients treated with cognitive behavioral therapy. J Affect Disord 1997;43:

131–142
4. Klein D. NIMH collaborative research on treatment of depression [comment]. Arch Gen Psychiatry 1990;47:682–688
5. Rush AJ, Prien R. From scientific knowledge to the clinical practice of psychopharmacology: can the gap be bridged? Psychopharmacol Bull 1995;31:7–20
6. Nierenberg AA, Mulroy R. Declaration of treatment failures. Mod Probl Pharmacopsychiatry 1997;25:17–33
7. Soury D, Amsterdam J, de Montigny C, et al. Treatment resistant depression: methodological overview and operational criteria. Eur Neuropsychopharmacol 1999;9:83–91
8. Feinstein AR. Clinimetrics. New Haven, Conn: Yale University Press; 1987
9. Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry 1960;23:56–62
10. Hamilton M. Development of a rating scale for primary depressive illness. Br J Soc Clin Psychol 1967;6:278–296
11. Williams JB. A structured interview guide for the Hamilton Depression Rating Scale. Arch Gen Psychiatry 1988;45:742–747
12. Carroll BJ, Fielding JM, Blashki TG. Depression rating scales: a critical view. Arch Gen Psychiatry 1973;28:361–366
13. Montgomery SA, Asberg MC. A new depression scale designed to be sensitive to change. Br J Psychiatry 1979;134:382–389
14. Rush AJ, Giles DE, Schlesser MA, et al. The Inventory for Depressive Symptomatology (IDS): preliminary findings. Psychiatry Res 1986;18: 65–87
15. Rush AJ, Gullion CM, Basco MR, et al. The Inventory for Depressive Symptomatology (IDS): psychometric properties. Psychol Med 1996;26: 477–486
16. Bech P, Gram LF, Dein E, et al. Quantitative rating of depressive states. Acta Psychiatr Scand 1975;51:161–170
17. Nierenberg AA, Keefe BR, Leslie VC, et al. Residual symptoms in depressed patients who respond acutely to fluoxetine. J Clin Psychiatry 1999;60:221–225
18. Rush AJ, Trivedi MH. Treating depression to remission. Psychiatr Ann 1995;25:704–705, 709
19. Guy W. ECDEU Assessment Manual for Psychopharmacology. US Dept Health, Education, and Welfare publication (ADM) 76-338. Rockville, Md: National Institute of Mental Health; 1976:218–222
20. Nierenberg AA, Wright EC. Evolution of remission as the new standard in the treatment of depression. J Clin Psychiatry 1999;60(suppl 22):7–11
21. National Institute of Mental Health. Sequenced Treatment Alternatives to Relieve Depression. Available at: http://www.edc.gsph.pitt.edu/stard. Accessed Feb. 23, 2001
22. Rush AJ, Thase ME. Psychotherapies for depressive disorders. In: Maj M, Sartorius N, eds. World Psychiatric Association Series on Evidence and Practice in Psychiatry, vol 1. Depressive Disorders. Chichester, England: John Wiley; 1999:161–206
23. Keller MB, McCullough JP, Klein DN, et al. A comparison of nefazodone, the cognitive behavioral-analysis system of psychotherapy, and their combination for the treatment of chronic depression. N Engl J Med 2000;342: 1462–1470
24. Eaves G, Rush AJ. Cognitive patterns in symptomatic and remitted unipolar major depression. J Abnorm Psychol 1984;93:31–40
25. Fava GA, Grandi S, Zielezny M, et al. Cognitive behavioral treatment of residual symptoms in primary major depressive disorder. Am J Psychiatry 1994;151:1295–1299
26. Paykel ES, Ramana R, Cooper Z, et al. Residual symptoms after partial remission: an important outcome in depression. Psychol Med 1995;25: 1171–1180
27. Mintz J, Mintz LI, Arruda MJ, et al. Treatments of depression and the functional capacity to work. Arch Gen Psychiatry 1992;49:761–768
28. Power MJ, Duggan CF, Lee AS, et al. Dysfunctional attitudes in depressed and recovered depressed patients and their first-degree relatives. Psychol Med 1995;25:1171–1180
29. Judd LL, Akiskal HS, Maser JD, et al. Major depressive disorder: a prospective study of residual subthreshold depressive symptoms as predictor of rapid relapse. J Affect Disord 1998;50:97–108
30. Thase ME, Simons AD, McGeary J, et al. Relapse after cognitive behavioral therapy of depression: potential implications for longer courses of treatment. Am J Psychiatry 1992;149:1046–1052
31. Keller MB, Lavori PW, Mueller TI, et al. Time to recovery, chronicity, and levels of psychopathology in major depression: a 5-year prospective follow-up of 431 subjects. Arch Gen Psychiatry 1992;49:809–816
32. Lavori PW, Keller MB, Scheftner W, et al. Recurrence after recovery in unipolar MDD: an observational follow-up study of clinical predictors and somatic treatment as a mediating factor. Int J Meth Psychiatr Res 1994;4: 211–220
33. Doogan DP, Caillard V. Sertraline in the prevention of depression. Br J Psychiatry 1992;160:217–222
34. Frank E, Kupfer DJ, Perel JM. Three-year outcomes for maintenance therapies in recurrent depression. Arch Gen Psychiatry 1990;47:1093–1099
35. Montgomery SA, Dufour H, Brion S. The prophylactic efficacy of fluoxetine in unipolar depression. Br J Psychiatry 1988;153(suppl 3):69–73
36. Montgomery SA, Dunbar G. Paroxetine is better than placebo in relapse prevention and prophylaxis of recurrent depression. Int Clin Psychopharmacol 1993;8:189–195
37. Montgomery SA, Rasmussen JGC, Tanghoj P. A 24-week study of 20 mg citalopram, 40 mg citalopram, and placebo in the prevention of relapse of major depression. Int Clin Psychopharmacol 1993;8:181–188
38. Robinson DS, Lerfald SC, Bennett B, et al. Continuation and maintenance treatment of major depression with the monoamine oxidase inhibitor phenelzine: a double-blind placebo-controlled discontinuation study. Psychopharmacol Bull 1991;27:31–39
39. Peselow ED, Dunner DL, Fieve RR, et al. The prophylactic efficacy of tricyclic antidepressants: a five-year follow-up. Prog Neuropsychopharmacol Biol Psychiatry 1991;15:71–82
40. Ramana R, Paykel E, Cooper Z. Remission and relapse in major depression: a two-year prospective follow-up study. Psychol Med 1995;25: 1161–1170
41. Nierenberg AA, Alpert JE. Depressive breakthrough. Psychiatr Clin North Am 2000;23:731–742