50% Improvement: Should Treatment Response Go Beyond Symptom Improvement When Evaluating the Treatment of Depression?

Mark Zimmerman, MD[a,*], and Sin-Ying Lin, PhD[a]

ABSTRACT

Background: The emphasis on symptom resolution in depression treatment research is at variance with the recommendations of official treatment guidelines and the results of surveys of depressed patients' views of the most important treatment goals. In the present study, we examined the interrelationship between response rates on various outcome domains and whether response on each domain was associated with patients' global rating of improvement (PGI) reported upon treatment completion. We also examined whether the PGI was associated with the number of domains on which the patients had achieved responder status and which domains were independent predictors of PGI response.

Methods: Between January 2016 to April 2022, 844 patients with *DSM-IV* major depressive disorder completed the Remission from Depression Questionnaire (RDQ), a self-report measure that assesses 6 constructs considered by patients to be relevant to assessing treatment outcome. The patients completed the RDQ at admission and discharge from the treatment program. For each domain, response was defined as a 50% or greater reduction in scores. At discharge, the patients rated the PGI.

Results: The patients significantly improved from admission to discharge on each of the 6 domains assessed on the RDQ (Cohen *d* range, 1.09–1.55). The responders on each domain reported significantly greater improvement on the global rating of improvement at discharge (all *P* values < .001). Responder status in one domain mostly co-occurred with responder status in another domain. In a logistic regression analysis, responses on all domains except nondepressive symptoms were independently associated with PGI response.

Conclusions: The results of the present study are consistent with the results of multiple patient surveys which have suggested that focusing on symptom reduction is too narrow of an approach when measuring outcome in the treatment of depression. Expanding the assessment of outcome beyond symptoms and viewing nonsymptomatic outcome domains as critical composites of primary endpoints would be more consistent with a patient-centered approach toward the treatment of depression.

J Clin Psychiatry 2023;84(3):22m14706

To cite: Zimmerman M, Lin S-Y. 50% improvement: should treatment response go beyond symptom improvement when evaluating the treatment of depression? *J Clin Psychiatry.* 2023;84(3):22m14706.

To share: https://doi.org/10.4088/JCP.22m14706

[a]Department of Psychiatry and Human Behavior, Brown Medical School, and the Department of Psychiatry, Rhode Island Hospital, Providence, Rhode Island

**Corresponding author:* Mark Zimmerman, MD, 146 West River St, 11B, Providence, RI 02904 (mzimmerman@lifespan.org).

The United States Food and Drug Administration's guidelines for evaluating the efficacy of antidepressant medication require only that medications reduce symptom severity[1,2]; other potential therapeutic objectives such as improvement in functioning are considered secondary endpoints. Accordingly, the primary outcome measure in most studies of the treatment of depression is a measure of symptom severity. Studies that report rates of treatment response or remission make such determinations based on symptom severity scales without consideration of other clinically important factors such as improvement in functioning and quality of life. Almost all pooled analyses, meta-analyses, and network analyses of depression treatments are based on measures of depression symptom severity.[3–6] Consistent with this, power analyses to determine sample size in controlled studies of depression treatments are based on estimates of improvement on depression symptom scales.

The narrow, symptom-focused approach toward defining outcome diverges from depressed patients' opinions regarding the most important factors to consider when evaluating treatment success. Our clinical research group asked patients to rate the importance of 16 factors in determining whether a depressive episode was in remission.[7] The patients judged a return to normal functioning, quality of life, the presence of positive aspects of mental health, general well-being, and the ability to cope with stress, in addition to symptom resolution, as most important in determining remission status.[7] Demyttenaere and colleagues[8] asked patients and clinicians to rate the importance of a range of items to be cured from depression and found that patients prioritized a restoration of positive affect, whereas clinicians emphasized resolution of depressive symptoms. An online international study of more than 2,000 depressed patients taking an antidepressant[9] found that improvement in work, family, or social functioning was almost twice as likely as improvement in mood to be identified as the primary treatment goal. A study of depressed patients receiving cognitive-behavioral therapy[10] found that interpersonal goals and coping with problems were 3 to 4 times more frequently endorsed than symptom goals. Other studies[11,12] have similarly found that many depressed patients give precedence to non-symptom goals.

The research suggesting that outcome in the treatment of depression should not be limited to symptom improvement led our clinical-research group to develop the Remission

Clinical Points

- The primary outcome measure in most studies of the treatment of depression is a measure of symptom severity, with response defined as a 50% or greater reduction in scores.
- Depressed patients often do not view symptom improvement as the most important goal of treatment but instead prioritize nonsymptom outcomes such as improvement in functioning, coping ability, and quality of life.
- The present study is the first to demonstrate that patients' perception of the overall benefit of a treatment intervention was significantly associated with responder status of both nonsymptom and symptom domains, thereby suggesting that focusing on symptom reduction is too narrow of an approach when measuring outcome in the treatment of depression.

from Depression Questionnaire (RDQ). The RDQ includes 6 subscales—depressive symptoms, nondepressive symptoms, coping ability, positive mental health, functional impairment, and quality of life/life satisfaction.[13] Because our research approach was patient-centered, the goal of our first study of the RDQ was to ascertain patients' opinion of the scale compared to a symptom scale—the Quick Inventory of Depressive Symptomatology (QIDS).[14] Approximately twice as many patients indicated that the RDQ, compared to the QIDS, allowed them to more accurately describe their current status, reflect the effectiveness of treatment, and better evaluate their goals of treatment.[15] Of note, twice as many patients expressed a preference to complete the RDQ to monitor their progress in treatment.

Subsequent studies established the reliability and validity of the RDQ. In two separate studies,[13,16] scores on the RDQ and QIDS were both significantly associated with patients' self-reported remission status. Moreover, after controlling for the QIDS scores, RDQ total scores remained significantly associated with remission status, whereas QIDS scores were not associated with remission status after controlling for RDQ scores. This suggested that the RDQ assesses constructs other than depressive symptoms that patients consider important in determining remission, a result that was consistent with our prior studies indicating that patients' perspectives of remission go beyond simply symptom resolution.

The multiple goals of the treatment of depression—symptom reduction, improved functioning, enhanced positive mental health, resilient coping, improved quality of life and life satisfaction—are not independent of each other. However, while these variables are correlated with one another,[17] they are not perfectly correlated with each other, and this is why it is clinically important to consider each of the factors when evaluating outcome and to understand the patients' perspective in their goals of treatment. For example, symptom improvement that is sufficient to enable the patient to return to work is a clinically significant level of improvement even if the change in scores on a depression symptom severity scale does not meet the threshold used to define a treatment response or remission.

A limitation of prior studies advocating a patient-centered approach toward outcome assessment in the treatment of depression is that the studies either were based on a cross-sectional design or were surveys of patients' opinions of what was important. Also, the analyses focused on the narrow concept of remission rather than a broader assessment of improvement in treatment.

In the current report from the Rhode Island Methods to Improve Diagnostic Assessment and Services (MIDAS) project, we examined treatment response to a brief, intensive intervention on each of the domains assessed by the RDQ as well as patients' global rating of improvement (PGI) on completion of treatment. Per standard practice in research reports, response was defined as a 50% or greater improvement in scores from the beginning to the end of treatment. We examined the interrelationship between response rates on the various domains and whether response on each domain was associated with PGI upon treatment completion, thus providing insight into the factors patients consider most important when evaluating treatment effectiveness. Further, we examined whether PGI was associated with the number of domains on which the patients had achieved responder status and which domains were independent predictors of patients' global rating of improvement. On the basis of our prior findings, we hypothesized that patients' ratings of improvement would be significantly associated with response in each of the RDQ domains and more highly associated with response in nonsymptom domains than in the depressive symptom domain.

METHODS

Patients

The study was conducted in the Rhode Island Hospital Department of Psychiatry partial hospital program, a 5-day/week intensive treatment program. Patients meet with a psychiatrist and therapist daily and attend 4 groups per day. The mean (SD) length of stay is 7.5 (4.7) days. A total of 844 patients with a principal diagnosis of major depressive disorder (MDD) who were evaluated at admission and discharge between January 2016 and April 2022 were included in the analysis. Patients who were admitted multiple times during the duration of the study had the data only from their first admission included in the data analysis. The Rhode Island Hospital institutional review committee approved the research protocol, and all patients provided informed, written consent.

Assessment

Almost half of the patients (44.43%, n = 375) were interviewed by a diagnostic rater who administered the Structured Clinical Interview for *DSM-IV* (SCID)[18] and the borderline personality disorder section of the Structured

Table 1. Admission and Discharge Scores and Response Rates on the RDQ Domains in 844 Depressed Patients[a]

RDQ Domain	Admission Score, Mean (SD)	Discharge Score, Mean (SD)	Cohen *d*	Responders, % (n) [95% CI]
Depressive symptom severity	19.2 (4.2)	11.1 (6.1)	1.55	38.7% (327) [35.4%–42.0%]
Nondepressive symptom severity	14.3 (4.8)	8.5 (5.5)	1.13	40.3% (340) [37.0%–43.6%]
Coping ability	7.4 (2.0)	4.6 (2.5)	1.18	30.6% (258) [27.0%–33.2%]
Positive mental health	19.2 (4.2)	11.8 (6.4)	1.37	33.6% (284) [29.9%–36.3%]
Functioning	13.0 (3.8)	8.2 (4.8)	1.09	33.6% (284) [29.6%–36.0%]
Well-being/life satisfaction	11.6 (2.7)	7.3 (4.0)	1.25	30.6% (258) [26.9%–33.1%]

[a]For each RDQ domain, response was defined as a 50% or greater improvement from admission to discharge.
Abbreviation: RDQ = Remission from Depression Questionnaire.

Interview for *DSM-IV* Personality (SIDP-IV).[19] Not all patients were evaluated with the semistructured interviews because of a lack of available interviewers. These patients were evaluated by board-certified psychiatrists.

The RDQ was the primary outcome measure in our program. We modified the RDQ to accommodate use with patients with varied diagnoses as seen in routine clinical practice as well as patients with multiple psychiatric disorders. Nineteen items were added to the original 41-item scale. The modified 60-item measure included 14 depressive symptoms, 11 nondepressive symptoms, 5 coping ability/stress tolerance items (eg, "I easily got overwhelmed by stress"), 12 positive mental health items (eg, "I felt confident"), 10 functioning items (eg, "I did not do my work [at a paid job, at home, or at school] as well as usual"), and 8 general well-being/life-satisfaction items (eg, "I was satisfied in my relationships"). The time frame is the past week, and the items are rated on a 3-point rating scale; the items were scored 0 (not at all or rarely true), 1 (sometimes true), and 2 (often or almost always true), with higher item values reflecting greater pathology and some items being reverse-scored. Thus, higher scores indicated greater symptomatology, poorer coping, more impaired functioning, fewer positive mental health indicators, and less life satisfaction. The reliability of the RDQ was previously studied in 274 depressed outpatients.[13] The scale had excellent internal consistency (Cronbach α = .97 for the total scale and above 0.80 for each of the subscales) and test-retest reliability (total scale $r = 0.85$ and above 0.60 for each subscale).[13] In the present study, the scale had excellent internal consistency (admission: Cronbach α = .92 for the total scale and above 0.65 for each of the subscales; discharge: Cronbach α = .97 for the total scale and above 0.80 for each of the subscales).

At discharge, the patients completed a program evaluation form that included an item asking the patients to rate their current state compared to when they entered the program: "Compared to how you were feeling when you first started the program, at the time of ending do you feel: 0 = no better 1 = slightly better 2 = moderately better 3 = a lot better 4 = very much better?"

Table 2. Patient Global Rating of Improvement as a Function of the Number of Remission From Depression Questionnaire (RDQ) Domains on Which the Patient Responded in 844 Depressed Patients[a]

No. of RDQ Domains in Which Patient Was a Responder	n	Global Rating of Improvement, Mean (SD)	Adjacent 2-Group *t* Tests	
			t Value	*P* Value
0	322	2.2 (1.0)		
			−5.77	< .001
1	114	2.6 (0.7)		
			−3.52	.001
2	105	3.0 (0.7)		
			2.04	.043
3	65	3.2 (0.7)		
			0.45	.655
4	59	3.2 (0.9)		
			−2.54	.012
5	78	3.5 (0.7)		
			−2.5	.013
6	101	3.7 (0.4)		

[a]Overall ANOVA: $F_{6, 837} = 67.4$; $P < .001$. The values in the columns under "Adjacent 2-Group *t* Tests" are set so that the values fall between the respective rows for which the comparisons are made.
Abbreviations: ANOVA = analysis of variance, RDQ = Remission from Depression Questionnaire.

Data Analysis

For each RDQ subscale, we computed the response rate (50% or greater improvement from admission to discharge). A 1-way analysis of variance (ANOVA) was used to compare the mean PGI score as a function of the number of RDQ subscales on which patients were responders, and follow-up pairwise comparisons would be conducted only if the overall ANOVA was significant. Patients with a PGI rating of 3 or 4 (a lot or very much better) were considered PGI responders. Chi-square tests were used to compare the percentage of PGI responders as a function of the number of RDQ domains on which the patients were responders. After the univariate analyses, we conducted a logistic regression analysis to determine on which outcome domains responder status was independently associated with PGI response.

Table 3. Mean PGI Score and PGI Response Rate in Patients Who Were and Were Not Responders on Each Domain of the RDQ in 844 Depressed Patients

RDQ Domain	PGI Score, Mean (SD)[a]				PGI Response Status			
	Responder	Nonresponder	*t* Value	*P* Value	Responder, % (n)	Nonresponder, % (n)	χ^2 Value	*P* Value
Depressive symptoms	3.4 (0.7)	2.4 (1.0)	15.92	<.001	91.4% (327)	52.2% (517)	138.44	<.001
Nondepressive symptoms	3.2 (0.8)	2.5 (1.0)	11.98	<.001	84.7% (340)	55.8% (504)	76.16	<.001
Coping ability	3.4 (0.7)	2.5 (1.0)	13.65	<.001	91.1% (258)	57.0% (586)	93.22	<.001
Positive mental health	3.5 (0.7)	2.4 (1.0)	17.53	<.001	94.7% (284)	53.6% (560)	143.37	<.001
Impaired functioning	3.4 (0.8)	2.5 (1.0)	14.47	<.001	91.2% (284)	55.4% (560)	108.56	<.001
Well-being/life satisfaction	3.5 (0.7)	2.5 (1.0)	18.22	<.001	95.7% (258)	55.0% (586)	133.82	<.001

[a]The global rating of improvement score was significantly higher in the responders on each RDQ subscale than the nonresponders on the RDQ subscale.
Abbreviations: PGI = patient global rating of improvement, RDQ = Remission from Depression Questionnaire.

Table 4. Concordance of Responder Status on the RDQ Domains in 844 Depressed Patients[a]

RDQ Domain	Depressive Symptoms (n = 327)	Nondepressive Symptoms (n = 340)	Coping Ability (n = 258)	Positive Mental Health (n = 284)	Impaired Functioning (n = 284)	Life Satisfaction (n = 258)
Depressive symptoms	…	71.5% (66.7%–76.3%)	72.9% (64.5%–78.3%)	76.4% (71.5%–81.3%)	71.5% (66.3%–76.8%)	77.9% (72.8%–83.0%)
Nondepressive symptoms	74.3% (69.6%–79.0%)	…	71.3% (65.8%–76.8%)	68.0% (62.6%–73.4%)	69.0% (63.6%–74.4%)	71.7% (66.2%–77.2%)
Coping ability	57.5% (52.1%–62.9%)	54.1% (48.8%–59.4%)	…	62.7% (57.1%–68.3%)	63.7% (58.1%–69.3%)	67.4% (61.7%–73.1%)
Positive mental health	66.4% (61.3%–71.5%)	56.8% (51.5%–62.1%)	69.0% (63.4%–74.6%)	…	70.1% (64.8%–75.4%)	82.6% (78.0%–87.2%)
Impaired functioning	62.1% (56.8%–67.4%)	57.6% (52.3%–62.9%)	70.2% (64.6%–75.8%)	70.1% (64.8%–75.4%)	…	75.2% (69.9%–80.5%)
Well-being/life satisfaction	61.5% (56.2%–66.8%)	54.4% (49.1%–59.7%)	67.4% (61.7%–73.1%)	75.0% (70.0%–80.0%)	68.3% (62.9%–73.7%)	…
Column mean	64.3% (59.1%–69.5%)	58.9% (53.7%–64.1%)	70.2% (64.6%–75.8%)	70.4% (65.1%–75.8%)	68.5% (63.1%–73.9%)	75.0% (69.7%–80.3%)

[a]Values are shown as percentage of responders (95% CI). The logic of the table presentation is as follows: there were 327 responders on the depressive symptoms subscale, and 74.3% were also responders on the nondepressive symptoms subscale, 57.5% were responders on the coping ability subscale, and so on. For each RDQ subscale, response was defined as a 50% or greater improvement from admission to discharge.
Abbreviation: RDQ = Remission from Depression Questionnaire.

Table 5. Logistic Regression Model of Responders on RDQ Domains Identifying PGI Response

Variable	Log Odds	SE	*z*	*P*
Intercept	−0.31	0.10	−3.01	.003
Depressive symptoms	1.09	0.26	4.24	<.001
Nondepressive symptoms	0.21	0.22	0.95	.341
Coping	0.58	0.28	2.08	.037
Positive mental health	1.24	0.33	3.79	<.001
Functioning	0.60	0.27	2.17	.030
Well-being	1.10	0.38	2.88	.004

Abbreviations: PGI = patient global rating of improvement, RDQ = Remission from Depression Questionnaire.

RESULTS

Demographic Characteristics and Response in Each Outcome Domain

The 844 patients included 550 (65.2%) cisgender female, 265 (31.4%) cisgender male, and 18 (2.1%) gender diverse individuals. Data on gender were missing for 11 patients. The patients ranged in age from 18 to 82 years (mean [SD] = 36.8 [13.9] years). Approximately half of the patients were single (45.7%, n = 386); the remainder were married (25.4%, n = 214), living with someone as if in a marital relationship (12.4%, n = 105), divorced (11.3%, n = 95), separated (2.7%, n = 23), or widowed (2.5%, n = 21). Over one-third of the patients completed a 4-year university degree (37.6%, n = 317). The majority of the sample identified as White (76.9%, n = 649). A minority of patients identified as Black (5.1%, n = 43), Hispanic (8.3%, n = 70), Asian (3.2%, n = 27), or from another or a combination of racial/ethnic backgrounds (6.5%, n = 55).

The patients showed significant levels of improvement from admission to discharge in all 6 domains (Table 1). The depressive symptoms domain had the largest effect size, and the functioning domain had the smallest effect size. The response rates on the 6 domains ranged from 30% to 40% (mean = 34.2%) (Table 1).

Number of Domains of Response and Patient Global Rating of Improvement

Nearly 40% of the patients were not responders on any of the 6 domains, and slightly more than 10% of the patients were responders on all 6 domains (Table 2). The patients were responders on a mean (SD) of 2.1 (2.2) domains.

The mean (SD) rating on the PGI was 2.8 (1.0), and 67.4% (n = 569) were PGI responders. PGI ratings increased as a function of the number of domains the patient was a responder ($F_{6, 837} = 67.4$, $P < .001$) (Table 2). Follow-up t tests comparing adjacent levels found that each 2-group comparison was significant except the difference between being a responder on 3 and 4 domains (Table 2).

Patient Global Rating of Improvement and Specific Domains of Response

For each outcome domain, we compared the responders and nonresponders on the mean PGI rating and the percentage of PGI responders. For both sets of analyses, the responders on each domain reported significantly greater improvement (Table 3).

Responder status in one domain was not independent of responder status in another domain (Table 4). Across all domains, the mean concordance of response between 2 domains was 67.9%. We also computed the phi correlations of agreement, and the mean of the phi coefficients was 0.51. The likelihood of a response in another domain was greatest for the responders in the life satisfaction domain (mean = 75.0%, mean phi coefficient = 0.55) and lowest for the responders in the nondepressive symptoms domain (mean = 58.9%, mean phi coefficient = 0.44). Responders in the depressive symptoms domain were least likely to respond on the coping ability domain (57.5%, phi = 0.46) and most likely to respond on the nondepressive symptom domain (74.3%, phi = 0.55) (Table 4).

Because of the concordance of response rates on the 6 domains, we further examined which domains independently predicted PGI response by conducting a logistic regression analysis. Response on all domains except nondepressive symptoms was independently associated with PGI response, with positive mental health as the most prominent predictor (Table 5). The total amount of variance accounted for by the entire set of variables was 25%.

Finally, because treatment studies of depression typically define response in terms of depressive symptoms, we repeated the logistic regression analysis entering the depression symptom domain first. The depression symptoms domain accounted for 15% of the variance of PGI response. The remaining variables accounted for another 10% of the variance. When we reversed this analysis and entered the nonsymptom domains first, they accounted for 22% of the variance of PGI response, and the symptom domains accounted for an additional 3% of the variance.

DISCUSSION

The emphasis on symptom resolution in research on the treatment of depression is at variance with the recommendations of official treatment guidelines of professional societies[20,21,22] which indicate that the overall goals of the treatment of depression are symptom resolution and reduced morbidity. In fact, some researchers suggest that improvement in functioning and/or quality of life should be the primary goal of treatment.[23]

Clinicians and patients have different perspectives on the primary goal of depression treatment. Whether treated with medication or psychotherapy, depressed patients often do not view symptom improvement as the most important goal of treatment. In the aforementioned survey of more than 2,000 patients taking an antidepressant in which nearly twice as many patients indicated that their primary goal of treatment was to improve functioning rather than mood,[9] it was also found that providers more frequently indicated that improvement in mood was the primary goal of treatment. The Depression and Bipolar Alliance surveyed the wellness priorities of more than 6,000 of their members and conducted a focus group of a small subset of the respondents.[24] Managing symptoms and not being controlled by symptoms were among the most frequently endorsed wellness priorities, and a theme emerging from the focus group participants was that one can experience symptoms and still be well, whereas providers overly focus on minimizing symptoms rather than improving quality of life. In another study comparing the themes arising from focus groups of patients and providers,[25] only patients emphasized improved functioning, managing depression, and acceptance of depression.

By contrast, when reviews, commentaries, and introductory sections of research studies quote statistics on the efficacy of treatment of depression, the percentages of responders, remitters, or treatment resisters are based solely on changes in scores on symptom severity scales. There is thus a disconnect between what patients value most and what researchers prioritize. We concur with Cuijpers'[26] recommendation that the perspective of patients should be given greater priority in studies of the efficacy and effectiveness of treatments for depression.

The current study goes beyond patient surveys of the goals of treatment and, to our knowledge, is the first to contemporaneously evaluate and demonstrate that patients' perception of the overall benefit of a treatment intervention was significantly associated with responder status of multiple outcome domains and that, after accounting for concordance in response on these domains, response on nonsymptom domains remained significant predictors of outcome.

The results of the study thus highlight the importance of broadening the assessment of outcome in treatment studies of depression. Improvement in functioning, well-being, positive mental health, and the ability to cope with stress should not be considered secondary endpoints but should be considered as important as improvement in symptoms. In fact, from depressed patients' perspective, these endpoints also could be considered as primary endpoints. Such an approach would move the field closer to what patients consider to be the sole or most important determinant of successful treatment. We doubt, however, that such a paradigm shift will occur until regulatory agencies that approve treatments endorse such a change.

While our discussion has implicitly focused on biological interventions because of our reference to regulatory agencies, it should be acknowledged that the emphasis on symptom reduction in treatment studies of depression is not limited to pharmacologic treatments. Granted, psychotherapies are not approved by a regulatory agency and thus are not required to demonstrate a reduction in symptom severity. However, the predominant outcome variable in psychotherapy studies of depression is also improvement in symptom levels.[26]

To be sure, symptom reduction is important. However, patients often link symptom improvement to improvement in other domains. For example, the desire to sleep better is linked to improved energy, which is connected to improved capability of fulfilling one's daily responsibilities. Symptom improvement in the absence of improved functioning, quality of life, and ability to cope with the daily stresses of life is a pyrrhic victory, analogous to an antihypertensive medication's reducing blood pressure but not reducing the risk of cardiovascular events.

The present study was conducted in a partial hospital program, which presents certain advantages and disadvantages. The therapeutic intervention is relatively brief and well-defined, and external factors such as the occurrence of major life events are less likely to influence outcome. Treatment in the partial hospital lasts approximately 6 hours per day; thus, most patients who are employed do not work during their time in the program, most students take a break from their studies, and the demands of childcare and household responsibilities are often reduced. These factors could reduce the impact of treatment on this outcome domain.

In the univariate analyses, the effect size of treatment for the nonsymptom domains was large, albeit somewhat smaller than the effect size of the depressive symptom domain. Given the short duration of the treatment program, this finding suggests that nonsymptom domains can respond as rapidly to treatment as symptoms. Unfortunately, we did not administer the RDQ repeatedly during the course of treatment, and thus we are unable to draw more definitive conclusions regarding the sequencing of improvement.

Almost all antidepressant efficacy trials exclude patients with some comorbid psychiatric disorders, and some trials exclude individuals with any comorbid disorder.[27] We did not exclude patients with comorbid disorders because most depressed patients have comorbid disorders,[28] and such an exclusion would reduce the generalizability/external validity of the study.

The results of the present study were based on self-administered questionnaires. It would be of interest to examine clinicians' global ratings of improvement and determine if the same factors underlying patients' improvement ratings also contribute to clinician ratings. Surveys and focus groups of clinicians suggest that clinicians' global ratings of improvement would be more heavily influenced by symptom improvement.[8,9,25]

A limitation of the study was that only a minority of the patients were evaluated with semistructured diagnostic interviews. We did not assess the reliability of clinicians' diagnoses, which is the usual case in real-world effectiveness studies.

The study was conducted in a single clinical program in which the majority of the patients were white and female and had health insurance. Replication in samples with different demographic characteristics is warranted. Moreover, the study was conducted in a partial hospital program where patients typically present in greater distress than outpatient settings. Reduction in scores on a measure of depressive symptoms partially reflects a reduction in distress, just as it does in studies of the effectiveness of treating depression in outpatient settings. This might account for why the largest univariate effect size was found for the depressive symptom domain. It will be important to replicate the findings of the current study of partial hospital patients in an outpatient sample, for whom most treatment for depression occurs, to further evaluate the relative importance of different domains in evaluating treatment outcome.

The focus of this article has been on responder status defined as a 50% or greater improvement in scores in each domain. The 50% threshold has been traditionally used to define responders on symptom severity scales. There is less precedent for this 50% threshold to define response in other domains. Utilizing the same definition for all domains seemed to us the most equitable way of comparing their relative influence on patients' global rating of improvement.

Responder status is a dichotomous variable. As such, there is no accounting for levels of improvement falling below the 50% threshold. Improvement less than 50% is often clinically significant. In a future analysis, we will further explore the impact of lower levels of improvement on patients' assessment of the overall effectiveness of treatment.

In conclusion, when evaluating the effectiveness of depression treatment in both research studies and clinical practice, it is important to consider the opinion of patients. The results of the present study are consistent with findings of multiple patient surveys which have suggested that focusing on symptom reduction is too narrow of an approach when measuring outcome in the treatment of depression, and they are consistent with the Lancet-World Psychiatric Association Commission on depression, which encouraged clinicians to elicit patients values and expectations regarding treatment goals.[29] Expanding the assessment of outcome beyond symptoms would be more consistent with a patient-centered approach toward the treatment of depression.

Submitted: October 26, 2022; accepted December 2, 2022.
Published online: May 8, 2023.
Relevant financial relationships: **Dr Zimmerman** has served on advisory boards for Intra-Cellular Therapies, Boehringer Ingelheim, GH Research, and Biogen. **Dr Lin** has no conflicts to disclose.
Potential conflicts of interest: None.
Funding/support: None.

REFERENCES

1. US Food and Drug Administration, Center for Biologics Evaluations and Research. Guidance for Industry: Providing clinical evidence of

effectiveness for human drugs and biological products. FDA website. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/providing-clinical-evidence-effectiveness-human-drug-and-biological-products. May 1998.
2. US Food and Drug Administration. Major depressive disorder: developing drugs for treatment. Guidance for industry Docket No. FDA-2018-D-1919. FDA website. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/major-depressive-disorder-developing-drugs-treatment. 2018:28851–28853.
3. Cipriani A, Furukawa TA, Salanti G, et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *Lancet*. 2009;373(9665):746–758.
4. Cuijpers P, Noma H, Karyotaki E, et al. A network meta-analysis of the effects of psychotherapies, pharmacotherapies and their combination in the treatment of adult depression. *World Psychiatry*. 2020;19(1):92–107.
5. Iovieno N, Papakostas GI. Correlation between different levels of placebo response rate and clinical trial outcome in major depressive disorder: a meta-analysis. *J Clin Psychiatry*. 2012;73(10):1300–1306.
6. Rutherford BR, Cooper TM, Persaud A, et al. Less is more in antidepressant clinical trials: a meta-analysis of the effect of visit frequency on treatment response and dropout. *J Clin Psychiatry*. 2013;74(7):703–715.
7. Zimmerman M, McGlinchey JB, Posternak MA, et al. How should remission from depression be defined? the depressed patient's perspective. *Am J Psychiatry*. 2006;163(1):148–150.
8. Demyttenaere K, Donneau AF, Albert A, et al. What is important in being cured from depression? discordance between physicians and patients (1). *J Affect Disord*. 2015;174:390–396.
9. Baune BT, Christensen MC. Differences in perceptions of major depressive disorder symptoms and treatment priorities between patients and health care providers across the acute, post-acute, and remission phases of depression. *Front Psychiatry*. 2019;10:335.
10. Holtforth MG, Wyss T, Schulte D, et al. Some like it specific: the difference between treatment goals of anxious and depressed patients. *Psychol Psychother*. 2009;82(Pt 3):279–290.
11. Ramnerö J, Jansson B. Treatment goals and their attainment: a structured approach to assessment and evaluation. *Cogn Behav Ther*. 2016;9:E2.
12. Uebelacker LA, Battle CL, Friedman MA, et al. The importance of interpersonal treatment goals for depressed inpatients. *J Nerv Ment Dis*. 2008;196(3):217–222.
13. Zimmerman M, Martinez JH, Attiullah N, et al. A new type of scale for determining remission from depression: the Remission from Depression Questionnaire. *J Psychiatr Res*. 2013;47(1):78–82.
14. 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.
15. Zimmerman M, Galione JN, Attiullah N, et al. Depressed patients' perspectives of two measures of outcome: The Quick Inventory of Depressive Symptomatology (QIDS) and the Remission from Depression Questionnaire (RDQ). *Ann Clin Psychiatry*. 2011;23(3):208–212.
16. Zimmerman M, Martinez JH, Attiullah N, et al. The remission from depression questionnaire as an outcome measure in the treatment of depression. *Depress Anxiety*. 2014;31(6):533–538.
17. Jones R, Yates WR, Williams S, et al. Outcome for adjustment disorder with depressed mood: comparison with other mood disorders. *J Affect Disord*. 1999;55(1):55–61.
18. First MB, Spitzer RL, Williams JBW, et al. *Structured Clinical Interview for DSM-IV (SCID)*. Washington, D.C.: American Psychiatric Association; 1997.
19. Pfohl B, Blum N, Zimmerman M. *Structured Interview for DSM-IV Personality*. Washington, DC: American Psychiatric Press, Inc.; 1997.
20. Gelenberg AJ, Freeman MP, Markowitz JC, et al. *American Psychiatric Association: Practice Guideline for the Treatment of Patients With Major Depressive Disorder*. 3rd ed. Washington, DC: American Psychiatric Association; 2010.
21. Malhi GS, Bassett D, Boyce P, et al. Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for mood disorders. *Aust N Z J Psychiatry*. 2015;49(12):1087–1206.
22. Lam RW, McIntosh D, Wang J, et al; CANMAT Depression Work Group. Canadian Network for Mood and Anxiety Treatments (CANMAT) 2016 clinical guidelines for the management of adults with major depressive disorder, Section 1: disease burden and principles of care. *Can J Psychiatry*. 2016;61(9):510–523.
23. IsHak WW, Greenberg JM, Balayan K, et al. Quality of life: the ultimate outcome measure of interventions in major depressive disorder. *Harv Rev Psychiatry*. 2011;19(5):229–239.
24. Morton E, Foxworth P, Dardess P, et al. "Supporting Wellness": a depression and bipolar support alliance mixed-methods investigation of lived experience perspectives and priorities for mood disorder treatment. *J Affect Disord*. 2022;299:575–584.
25. Kan K, Jörg F, Buskens E, et al. Patients' and clinicians' perspectives on relevant treatment outcomes in depression: qualitative study. *BJPsych Open*. 2020;6(3):e44.
26. Cuijpers P. Targets and outcomes of psychotherapies for mental disorders: an overview. *World Psychiatry*. 2019;18(3):276–285.
27. Zimmerman M, Clark HL, Multach MD, et al. Have treatment studies of depression become even less generalizable?: a review of the inclusion and exclusion criteria in placebo controlled antidepressant efficacy trials published during the past 20 years. *Mayo Clin Proc*. 2015;90(9):1180–1186.
28. Zimmerman M, Chelminski I, McDermut W. Major depressive disorder and Axis I diagnostic comorbidity. *J Clin Psychiatry*. 2002;63(3):187–193.
29. Herrman H, Patel V, Kieling C, et al. Time for united action on depression: a Lancet–World Psychiatric Association Commission. *Lancet*. 2022;399(10328):957–1022.