A Clinically Useful Anxiety Outcome Scale

Mark Zimmerman, MD; Iwona Chelminski, PhD;
Diane Young, PhD; and Kristy Dalrymple, PhD

Objective: Standardized scales are increasingly being recommended to measure outcome when treating psychiatric disorders in routine clinical practice. If the standard of care is to change and scales are to be incorporated into clinical practice, then it will be necessary to develop measures that are feasible to use as well as have good psychometric properties. In the present report from the Rhode Island Methods to Improve Diagnostic Assessment and Services project, we describe the reliability and validity of the Clinically Useful Anxiety Outcome Scale (CUXOS). The CUXOS was designed to be a brief (completed in less than 2 minutes), quickly scored (in less than 15 seconds), clinically useful measure that is reliable, valid, and sensitive to change.

Method: Nearly 1,000 psychiatric outpatients completed the CUXOS and were rated on clinician severity indices of depression, anxiety, and anger. A subset of patients completed other self-report symptom severity scales in order to examine discriminant and convergent validity, and a subset completed the CUXOS twice in order to examine test-retest reliability. Sensitivity to change was examined in patients with panic disorder and generalized anxiety disorder.

Results: On average, the CUXOS took less than 1.5 minutes to complete. The scale had high internal consistency and test-retest reliability, and was more highly correlated with other self-report measures of anxiety than with measures of depression, substance use problems, eating disorders, and anger. The CUXOS was more highly correlated with clinician severity ratings of anxiety than with depression and anger, and CUXOS scores were significantly higher in psychiatric outpatients with anxiety disorders than in patients with other psychiatric disorders. Finally, the CUXOS was a valid measure of symptom change.

Conclusions: The results of this large validation study of the CUXOS show that it is a reliable and valid measure of anxiety that is feasible to incorporate into routine clinical practice.

J Clin Psychiatry 2010;71(5):534–542

Submitted: *April 6, 2009; accepted November 10, 2009.*
Online ahead of print: *March 9, 2010 (doi:10.4088/JCP.09m05264blu).*
Corresponding author: *Mark Zimmerman, MD, Bayside Medical Center, 235 Plain St, Providence, RI 02905 (mzimmerman@lifespan.org).*

To determine the impact of treatment of any medical disorder it is necessary to evaluate outcome. Like other health care providers, mental health clinicians routinely evaluate outcome in their patients. However, in clinical practice these outcome assessments are typically based on unstructured interactions, thereby yielding global, impressionistic, judgments of progress rather than a quantified index of outcome. This is at variance with other areas of medical care in which outcome is determined, in part, on the change of a numerical value. Body temperature, blood pressure, cholesterol values, blood sugar levels, cardiac ejection fraction, and white blood cell counts are examples of quantifiable variables that are used to evaluate treatment progress. In the mental health field, standardized, quantifiable outcome measures exist for most major psychiatric disorders, yet they are rarely routinely incorporated in routine clinical practice. Gilbody and colleagues[1] surveyed 340 psychiatrists in the United Kingdom regarding their use of outcome measures. Only 11.2% of the psychiatrists routinely used standardized measures to assess outcome when treating depression and anxiety disorders. More than half of the clinicians indicated that they never used standardized measures to evaluate outcome. Zimmerman and McGlinchey[2] conducted a similar survey of 317 psychiatrists in the United States and found that less than 20% of the psychiatrists use scales to measure symptom severity at every visit or nearly every visit.

Outcome assessment is assuming increasing importance in this country, and while psychiatrists have not yet embraced the use of standardized scales in clinical practice, payor mandates may accelerate a change in clinicians' behavior. The Centers for Medicare and Medicaid Services' Physician Quality Reporting Initiative (PQRI),[3] signed into law in 2006, is intended to improve quality of care by providing physicians financial incentives to document outcomes reflecting best practices. In 2007, the first year of the PQRI, 74 indicators were listed, 1 of which was related to the treatment of depression. In 2008, the PQRI list of indicators was expanded to 134 items, with 2 additional indicators related to the treatment of depression. And in 2009 the PQRI list of indicators was expanded to 153 indicators, although none of the additions were related to the assessment or treatment of a psychiatric disorder. To date, no indicators have been related to anxiety, although a future version of the PQRI may well include indicators related to the recognition and management of clinically significant anxiety.

If the optimal delivery of mental health treatment ultimately depends on examining outcome, then precise, reliable, valid, informative, and user-friendly measurement is critical to evaluating the quality and efficiency of care in clinical practice. Clinicians are already overburdened with paperwork, and adding to this load by requiring repeated detailed evaluations with such instruments as the Hamilton Anxiety Rating Scale[4] is unlikely to meet with success. Self-report questionnaires are a cost-effective option because they are inexpensive in terms of professional time needed for administration, and they correlate highly with clinician ratings. Moreover, self-report scales are free of clinician bias and are therefore free from clinician overestimation of patient improvement (which might occur when there is an incentive to document treatment success).

In a discussion of our development of a clinically useful scale for assessing depression,[5] we identified 3 consumers who should be considered in the construction of a self-administered outcome questionnaire to be used in routine clinical practice: the patient, the clinician, and the administrator. Patients should find the measure user-friendly and the directions easy to follow. The questions should be understandable and relevant to the patient's problem. The scale should be brief, taking no more than 2 to 3 minutes to complete, so that upon routine administration at follow-up visits patients are not inconvenienced by the need to come for their appointment 10–15 minutes early in order to complete the measure. This would make it feasible to have the scale completed at each follow-up visit in the same way that blood pressure, body temperature, and weight are routinely assessed in primary care settings.

The instrument should provide clinicians with clinically useful information and improve the efficiency of conducting their clinical evaluation; thus, the measure should have practical value to the practicing clinician. Of course, clinicians need to be able to trust the information provided by any instrument they use. Consequently, outcome measures should have a sound basis in psychometrics, demonstrating good reliability, validity, and sensitivity to change. Clinicians and clinics should also find the instrument user-friendly; it should be easy to administer and score, with minimal training.

Clinic administrators likewise want measures to be both reliable and valid. To successfully implement an outcomes assessment program, administrators want measures to have high patient and clinician acceptance. Administrators are also concerned about the cost of an instrument, from the perspective of both the purchase price and the cost of labor to score the scale. Thus, an outcome measure, or outcome assessment program, should be inexpensive to purchase and implement.

Finally, we believe that any instrument constructed for use in clinical settings should meet scientific standards for publication in peer-reviewed journals. It is important that a new measure stand up to critical scientific review and be published in the scientific arena so that other investigators may further examine its properties.

During the past decade we have established and have been conducting the Rhode Island Methods to Improve Diagnostic Assessment and Services (MIDAS) project.[6,7] One of the goals of the MIDAS project has been to develop instruments for use in routine clinical practice. Previously we have described the reliability and validity of a broad-based self-report scale for psychiatric screening.[8–10] We have also developed 3 self-report depression scales that vary in respondent burden—from a single-item global measure of depression severity[11] to a case-finding measure that consists of approximately 100 statements and includes detailed assessments of psychosocial impairment and quality of life.[12] Intermediate between these extremes is the Clinically Useful Depression Outcome Scale (CUDOS)[5] that, similar to the Diagnostic Inventory for Depression,[12] also covers all of the *DSM-IV* diagnostic criteria for major depressive disorder but which uses Likert ratings of symptom statements in order to keep the scale brief enough to be feasibly incorporated into routine clinical practice.[5,13,14] It is the latter scale that we include in our clinical practice to evaluate outcome because it is brief enough so that respondent burden is minimal[15] yet also clinically useful because it evaluates each of the symptoms of major depression.

In the present report from the MIDAS project, we extend our work beyond depression and describe the reliability and validity of the Clinically Useful Anxiety Outcome Scale (CUXOS). Similar to the CUDOS, the CUXOS was designed to be brief (completed in less than 3 minutes), quickly scored (in less than 15 seconds), clinically useful, reliable, valid, and sensitive to change. Certainly, there is no shortage of self-report anxiety scales.[16] However, some are too long,[17–19] are expensive to purchase,[20,21] or are somewhat complicated to score.[22] These factors reduce their appeal as outcome tools for use in routine clinical practice. Other brief scales have also been developed, but most assess only a single anxiety disorder rather than the broader construct of anxiety.[23–25]

METHOD

Individuals presenting for an intake evaluation at the Rhode Island Hospital Department of Psychiatry outpatient practice were asked to complete the CUXOS as part of their initial paperwork. Because we were planning to test the CUXOS' validity by examining its relationship with psychiatric diagnoses, the diagnosticians were kept blind to the subjects' responses on the measure. The Rhode Island Hospital institutional review committee approved the research protocol, and all patients provided informed, written consent.

We conceptualized the CUXOS as a general measure of psychic and somatic anxiety rather than a disorder-

specific scale. We developed the scale in this manner so that it would be useful in the management of depressed patients, who often report high levels of anxiety in the absence of a specific anxiety disorder,[26,27] as well as be useful in the monitoring of patients with a variety of diagnosed anxiety disorders.

The content of the CUXOS items was derived from clinician rating scales such as the Hamilton Anxiety Rating Scale and *DSM-III-R* and *DSM-IV* descriptions of panic disorder and generalized anxiety disorder. The pool of items was reviewed by clinicians experienced in treating anxiety and mood disorders and revised accordingly. The initial version of the CUXOS included 25 items. The respondent is instructed to rate the CUXOS items on a 5-point Likert scale indicating "how well the item describes you during the past week, including today" (0 = not at all true; 1 = rarely true; 2 = sometimes true; 3 = usually true; 4 = almost always true).

All patients were interviewed by a trained diagnostic rater who administered the Structured Clinical Interview for *DSM-IV* (SCID),[28] supplemented with questions from the Schedule for Affective Disorders and Schizophrenia (SADS)[29] assessing the severity of symptoms during the week prior to the evaluation. One of the items on the SADS assesses psychic anxiety and another assesses somatic anxiety. We examined the association between the CUXOS and each of these ratings of anxiety as well as the combined total of the 2 items. Details regarding interviewer training and diagnostic reliability are available in other publications from the MIDAS project.[6,7,30]

To examine the convergent and discriminant validity of the CUXOS, we instructed subjects to complete a booklet of questionnaires at home that included measures of symptoms related to bulimia (Eating Disorder Inventory Anorexia and bulimia subscales[31]), depression (Beck Depression Inventory [BDI][32]), social phobia (Brief Fear of Negative Evaluation Scale[33]), (Fear Questionnaire-Social Phobia Subscale[34]), agoraphobic fears and cognitions (Fear Questionnaire-agoraphobia subscale[34]), (Social Phobia and Anxiety Inventory-agoraphobia subscale[19]) posttraumatic stress disorder (Posttraumatic Stress Disorder Scale[35]), obsessive-compulsive behavior (Maudsley Obsession-Compulsion Questionnaire[36]), cognitions common in generalized anxiety (Penn State Worry Scale[24]), anxiety symptoms and cognitions common in panic disorder (Beck Anxiety Inventory[20]), alcohol use (Michigan Alcohol Screening Test[37]), drug use (Drug Abuse Screening Test[38]), hypochondriasis (Whitely Index[39]), and somatization (Somatic Symptom Index[40,41]). These scales have been widely used, and their reliability and validity well established.

The test-retest reliability of the CUXOS was examined in 216 patients who completed the CUXOS at the time of their first appointment and were given the scale at the conclusion of the intake evaluation and asked to mail it back in a preaddressed postage-paid envelope. They were told that the purpose of the second administration was to test the performance of the scale, not to question the truthfulness or accuracy of their responses. Patients completed the second administration an average of 4.1 days (SD = 5.7) after the initial testing.

Sensitivity to change was examined in 43 patients with generalized anxiety disorder and 35 patients with panic disorder who completed the CUXOS within 6 months after treatment had begun (mean = 12.1 weeks, range = 4 to 22 weeks). At the follow-up evaluation, the treating clinicians rated the patients on the Clinical Global Impressions-Improvement scale (CGI-I)[42] and the Global Assessment of Functioning (GAF). The ratings at follow-up were made blind to the scores on the CUXOS.

Scale completion time was recorded in 58 depressed or anxious psychiatric outpatients at a follow-up visit with their psychiatrist. The group included 9 (15.5%) men and 49 (84.5%) women who ranged in age from 18 to 84 years (mean = 46.7, SD = 13.9).

Data Analyses

We undertook a sequence of 7 analyses. First, to reduce item redundancy, a correlation matrix of all items was generated to identify items that were highly correlated and thus could be eliminated from the scale. Second, an exploratory principal components analysis followed by varimax rotation was used to determine the initial factor structure and to determine if there were separate psychic and somatic anxiety factors. Factors with eigenvalues greater than 1 were retained. Third, we examined 2 types of reliability of the CUXOS—test-retest reliability and internal consistency. Fourth, we examined convergent and discriminant validity[43] by comparing the correlation between the CUXOS and measures of anxiety, such as the Beck Anxiety Inventory and Penn State Worry Questionnaire, with the correlation between the CUXOS and measures of depression, substance use, eating disorders, and somatization. Because anxiety disorders are frequently comorbid with other Axis I disorders, we predicted that the CUXOS would be significantly correlated with measures of other symptom domains, although the correlation with other measures of anxiety would be significantly higher. To further examine the convergent and discriminant validity of the CUXOS, we compared correlations between the scale and SADS clinical ratings of anxiety, depression, and anger. We tested whether the difference in the magnitude of the correlations was statistically significant by using Steiger's *z*, a test of the equality of 2 dependent correlations.[44] Fifth, to determine whether CUXOS scores were able to discriminate between levels of anxiety severity, we conducted an analysis of variance on the SADS anxiety severity ratings, followed by Tukey honestly significant difference (HSD) follow-up tests between each adjacent level of severity on the SADS ratings. Sixth, we used *t* tests to determine whether CUXOS scores were statistically significantly higher in patients with

specific anxiety disorders compared to psychiatric patients with other diagnoses. We used Levene's test for equality of variances to examine homogeneity of variance of the 2 samples, and, when statistically significant, used separate variance estimates with adjusted degrees of freedom. And seventh, we use *t* tests and paired *t* tests to examine the sensitivity of the CUXOS to change in patients with panic disorder and generalized anxiety disorder.

RESULTS

Demographic and Clinical Characteristics

The 963 patients in the study included 379 (39.4%) men and 584 (60.6%) women who ranged in age from 18 to 80 years (mean = 38.1, SD = 13.0). About two-fifths of the subjects were married (41.5%, n = 400); the remainder were single (30.0%, n = 289), divorced (13.5%, n = 130), separated (6.4%, n = 62), widowed (1.9%, n = 18), or living with someone as if in a marital relationship (6.6%, n = 64). The educational level achieved by the subjects was 10.0% (n = 96) did not graduate high school, 64.4% (n = 620) graduated high school or achieved equivalency, and 25.6% (n = 247) graduated college. The racial composition of the sample was 86.7% (n = 835) white, 4.3% (n = 41) black, 2.8% (n = 27) Hispanic, 0.8% (n = 8) Asian, and 5.3% (n = 52) from another or a combination of the above racial backgrounds.

The data in Table 1 show the diagnostic characteristics of the 963 patients who completed the CUXOS at their initial appointment. The most frequent *DSM-IV* diagnoses were major depressive disorder (44.5%), social phobia (28.2%), generalized anxiety disorder (17.1%), and panic disorder (18.1%). For the entire sample, the mean of the SADS psychic anxiety rating was 2.1 (SD = 1.5), and the average somatic item score was 1.8 (SD = 1.5). For comparison, the mean SADS depressed mood item score was 2.7 (SD = 2.6). Thus, as a group, the patients exhibited an equally mild-moderate level of depression and anxiety.

Table 1. Current *DSM-IV* Axis I Diagnoses of 963 Psychiatric Outpatients

DSM-IV Diagnosis[a]	n	%
Major depressive disorder	429	44.5
Bipolar disorder	49	5.1
Dysthymic disorder	73	7.6
Generalized anxiety disorder	165	17.1
Panic disorder	174	18.1
Social phobia	272	28.2
Specific phobia	91	9.4
Obsessive-compulsive disorder	73	7.6
Posttraumatic stress disorder	115	11.9
Adjustment disorder	54	5.6
Schizophrenia	4	0.4
Eating disorder	66	6.9
Alcohol abuse/dependence	80	8.3
Drug abuse/dependence	49	5.1
Somatoform disorder	60	6.2
Attention-deficit disorder	41	4.3
Impulse control disorder	41	4.3

[a]Individuals could be given more than 1 diagnosis.

Elimination of Redundant and Infrequently Occurring Items

Before examining the psychometric performance of the scale, we first examined the interitem matrix of correlations of the original 25 items on the scale. To reduce redundancy, we retained only 1 item of a pair that was similar in content and that correlated 0.75 or higher. In determining which item to retain, we examined each item's test-retest reliability, correlations with the anxiety ratings on the SADS, and the item-scale correlations. Four items were eliminated ("I trembled or felt shaky" [correlated 0.78 with the retained item: "I felt jittery"]; "I had difficulty relaxing" [correlated 0.76 with the retained item: "I felt keyed up and on edge"]; "I was unsteady on my feet" [correlated 0.76 with retained item: "I was dizzy or lightheaded"]; and "I felt terrified" [correlated 0.83 with the retained item: "I felt scared"]). We examined the frequency of the remaining 21 items in patients scoring 3 or more on either the SADS psychic or somatic anxiety items. Items that did not occur with at least 25% frequency in patients with moderate anxiety were considered too infrequent to be included in an outcome scale. This resulted in the exclusion of 1 item ("I had choking feelings.") This left a total of 20 items. The items are listed in Table 2. The subsequent psychometric analyses are based on the 20-item scale.

Scale Completion Time

After the above analyses were completed, and the final version of the CUXOS was drafted, we administered the measure to 58 psychiatric outpatients in ongoing treatment and recorded the time it took to complete the scale. All but 5 patients completed the scale in less than 2 minutes (mean = 78.4 seconds, SD = 35.0).

Factor Analysis, Internal Consistency and Test-Retest Reliability of the CUXOS

The mean total score on the CUXOS was 25.8 (SD = 19.5). The mean scores on the 6-item psychic anxiety subscale and 14-item somatic anxiety subscale were 10.6 (SD = 7.1) and 15.3 (SD = 13.7), respectively.

A factor analysis yielded 2 factors with eigenvalues greater than 1. The 2 factors accounted for 58.7% of the total variance. On the first factor, the 14 somatic anxiety items had the highest loadings, each item with a loading ≥ 0.49. On the second factor, the 6 psychic anxiety items had the highest loadings with each item loading ≥ 0.68.

The internal consistency of the CUXOS total scale score and the psychic and somatic anxiety subscales was examined in the entire sample who completed the scale at the initial evaluation and a sample of 120 patients who completed it at a follow-up appointment. The CUXOS demonstrated excellent internal consistency (baseline: Cronbach α = .95 for

Table 2. Item-Total Correlations and Test-Retest Reliability of Individual Clinically Useful Anxiety Outcome Scale (CUXOS) Items

CUXOS Item	Item-Total Correlations[a] Baseline (n = 963)	Item-Total Correlations[a] Follow-Up (n = 120)	Test-Retest Reliability[a] (n = 216)
I felt nervous or anxious	0.70	0.76	0.76
I worried a lot that something bad would happen	0.68	0.74	0.71
I worried too much about things	0.62	0.78	0.70
I was jumpy and easily startled by noises	0.65	0.77	0.80
I felt keyed up and on edge	0.72	0.80	0.74
I felt scared	0.69	0.76	0.77
I had muscle tension or muscle aches	0.65	0.62	0.74
I felt jittery	0.78	0.71	0.78
I was short of breath	0.73	0.73	0.80
My heart was pounding or racing	0.78	0.77	0.81
I had cold, clammy hands	0.73	0.69	0.74
I had a dry mouth	0.66	0.49	0.78
I was dizzy or lightheaded	0.72	0.66	0.75
I felt sick to my stomach (nauseated)	0.68	0.68	0.72
I had diarrhea	0.49	0.50	0.74
I had hot flashes or chills	0.71	0.65	0.78
I urinated frequently	0.54	0.50	0.70
I felt a lump in my throat	0.63	0.66	0.75
I was sweating	0.69	0.68	0.72
I had tingling feelings in my fingers or feet	0.63	0.56	0.77

[a]All correlations are significant at $P < .001$.

Table 3. Discriminant and Convergent Validity of the Clinically Useful Anxiety Outcome Scale (CUXOS)

Scale	Correlation With CUXOS, r[a]
Measure of anxiety	
Beck Anxiety Inventory	0.79
Penn State Worry Scale	0.54
Social Phobia and Agoraphobia Inventory—agoraphobia subscale	0.63
Anxiety Sensitivity Index	0.63
Posttraumatic Stress Disorder Scale	0.45
Maudsley Obsession—compulsive questionnaire	0.54
Fear Questionnaire social phobia subscale	0.41
Measure of nonanxious symptom domains	
Eating Disorder Inventory bulimia subscale	0.19
Eating Disorder Inventory anorexia subscale	0.21
Self-Report Mania Inventory	0.37
Symptom Rating Test paranoia subscale	0.40
Symptom Rating Test psychosis subscale	0.32
Beck Depression Inventory	0.55
Michigan Alcohol Screening Test	0.10
Drug Abuse Screening Test	0.18
State-Trait Anger Inventory-State Scale	0.36

[a]Because of missing data, the sample sizes ranged from 237 to 309. All correlations are significant at $P < .001$, except Michigan Alcohol Screening Test (nonsignificant) and Drug Abuse Screening Test ($P < .05$).

the total scale, .90 for the psychic anxiety subscale, and .93 for the somatic anxiety subscale; follow-up: α = .95 for the total scale, .94 for the psychic anxiety subscale, and .91 for the somatic anxiety subscale). The data in Table 2 show the correlation between each item and the total scale score. All item-scale correlations were statistically significant at baseline (median = 0.68) and at follow-up (median = 0.68).

The test-retest reliability of the CUXOS was examined in 216 subjects who completed the scale at the time of their initial appointment. The test-retest reliability of the total scale and the psychic and somatic anxiety subscales was high ($r = 0.90$, 0.86, 0.88, respectively), and the test-retest reliability of each item was statistically significant (median $r = 0.75$) (Table 2).

Discriminant and Convergent Validity of the CUXOS

Three hundred nine patients completed a package of questionnaires at home an average of 1.2 days (SD = 16.9) after the intake evaluation. The data in Table 3 show that the CUXOS was more highly correlated with measures of anxiety (median $r = 0.54$) than with measures of the other symptom domains (median $r = 0.32$). To be sure, the CUXOS was significantly correlated with measures of nonanxiety symptom domains because of the interrelationship between anxiety and other forms of psychopathology. The highest correlation with a nonanxiety measure was with the BDI ($r = 0.55$), although this was significantly less than the correlation between the CUXOS and the Beck Anxiety Inventory ($r = 0.79$).

To further explore the relationship between the CUXOS and ratings of affective dimensions, we examined the association between the scale and clinicians' rating on the SADS of depressed mood, psychic anxiety, and anger. The ratings were made blind to scores on the CUXOS. The correlation was highest with the anxiety rating ($r = 0.59$, $P < .001$), although the correlations with ratings of depression ($r = 0.37$, $P < .001$) and anger ($r = 0.27$, $P < .001$) were also significant. The correlation with the anxiety rating was significantly higher than the correlation with depression ($z = 7.4$, $P < .01$) and anger ($z = 9.6$, $P < .01$) ratings.

The Ability of the CUXOS to Discriminate Between Levels of Anxiety Severity

The ability of the CUXOS to discriminate between different levels of anxiety severity was examined with an analysis of variance on the SADS anxiety severity ratings. We separately examined the relationship between ratings on the SADS psychic anxiety item and the CUXOS psychic anxiety subscale, and the SADS somatic anxiety item and the CUXOS somatic anxiety subscale. The data in Table 4 show that increasing SADS severity ratings were associated with significantly higher CUXOS subscale scores. For the psychic anxiety subscale, Tukey HSD follow-up tests found that the difference between each adjacent level of severity (eg, nonanxious vs minimally anxious; mild vs moderate) was significant except for the comparison between nonanxious and minimally anxious patients ($P < .09$). For the somatic anxiety subscale, Tukey HSD follow-up tests found that

Table 4. SADS Psychic and Somatic Anxiety Severity Ratings and Mean Clinically Useful Anxiety Outcome Scale (CUXOS) Subscale Scores[a]

	CUXOS Psychic Anxiety Subscale		CUXOS Somatic Anxiety Subscale	
SADS Severity Rating	n	Mean (SD)	n	Mean (SD)
0 (none)	234	5.6 (6.0)	311	7.6 (10.1)
1 (minimal)	117	7.3 (5.4)	107	12.4 (10.6)
2 (mild)	181	9.9 (6.1)	211	15.5 (11.6)
3 (moderate)	230	13.0 (6.2)	208	20.9 (13.1)
4–5 (severe)	201	16.2 (5.7)	126	28.3 (14.5)

[a]SADS indicates Schedule for Affective Disorders and Schizophrenia. On the SADS, the severity of psychic and somatic anxiety are separately rated from 0 to 5. In the present analysis, the mean scores on the CUXOS psychic anxiety subscale is shown for the SADS psychic anxiety severity ratings, and the mean scores on the CUXOS somatic anxiety subscale correspond to the SADS somatic anxiety severity rating. The 5-group analysis of variance was significant in both analyses (psychic anxiety: $F_{4,958} = 104.3$, $P < .001$; somatic anxiety: $F_{4,958} = 83.8$, $P < .001$).

Abbreviation: SADS = Schedule for Affective Disorders and Schizophrenia.

the difference between each adjacent level of severity was statistically significant except for the comparison between minimally and mildly anxious patients.

Association With Psychiatric Diagnosis

Patients with any *DSM-IV* anxiety disorder (n = 556) scored significantly higher than patients with no current anxiety disorder (n = 407) (32.9 ± 19.8 vs 16.8 ± 14.6, $df = 961$, $t = 14.5$, $P < .001$). We examined CUXOS scores in patients with each of the *DSM-IV* anxiety disorders. The comparison group in each of these analyses was the 407 patients without a current anxiety disorder. The data in Table 5 show that for each anxiety disorder, patients with the disorder scored significantly higher than patients with no current anxiety disorder.

Because anxiety disorders frequently co-occur, some disorders may have been significantly associated with CUXOS scores by virtue of their association with other anxiety disorders. The majority of patients with each anxiety disorder were diagnosed with at least 1 other anxiety disorder (panic disorder, 70.7%; generalized anxiety disorder, 70.3%; social phobia, 65.8%; specific phobia, 75.8%; posttraumatic stress disorder, 70.4%; obsessive-compulsive disorder, 72.6%). We conducted a second series of analyses and included in the index group patients with the index anxiety disorder and no other anxiety disorder. For example, the panic disorder group included the 48 patients with panic disorder and no other anxiety disorder. In each of these analyses, the comparison group remained the 407 patients without a current anxiety disorder. The exclusion of patients with comorbid disorders reduced the size of the obsessive-compulsive disorder and specific phobia groups below 25; therefore, we did not compare them to the nonanxious group. For each of the remaining anxiety disorders, patients with the "pure" noncomorbid anxiety disorders scored significantly higher on the CUXOS (panic disorder: 39.6 ± 18.8 vs 16.8 ± 14.6, $df = 53.9$, $t = 8.1$, $P < .001$; generalized anxiety disorder: 26.9 ± 17.7 vs 16.8 ± 14.6, $df = 480$, $t = 5.3$, $P < .001$; social phobia: 22.2 ± 16.0 vs 16.8 ± 14.6, $df = 491$, $t = 3.0$, $P < .01$; posttraumatic stress disorder: 30.8 ± 19.6 vs 16.8 ± 14.6, $df = 32.6$, $t = 3.9$, $P < .001$).

Table 5. Clinically Useful Anxiety Outcome Scale (CUXOS) Scores in Psychiatric Outpatients With and Without a Current *DSM-IV* Anxiety Disorder

Current Anxiety Disorder	CUXOS Total Score, Mean (SD)	t[a]	P Value
Panic disorder	43.5 (19.4)	16.8	< .001
Generalized anxiety disorder	35.7 (19.3)	12.9	< .001
Social phobia	32.6 (20.0)	11.2	< .001
Specific phobia	35.2 (20.6)	8.1	< .001
Obsessive-compulsive disorder	33.4 (20.4)	6.7	< .001
Posttraumatic stress disorder	40.4 (21.8)	10.9	< .001
No anxiety disorder	16.8 (14.6)		

[a]CUXOS scores were compared between each anxiety disorder and the no anxiety disorder (n = 407) group.

Sensitivity to Change

Thirty-five patients with panic disorder completed the CUXOS 4–22 weeks (mean = 12.4, SD = 3.4) after initiating treatment. At follow-up, 17 (48.6%) patients were rated much or very much improved on the CGI-I, and mean ratings on the GAF were 60.5 (SD = 12.3). At the follow-up visit, the patients who were much or very much improved (ie, treatment responders) scored significantly lower than the patients who were not treatment responders. (16.7 ± 15.5 vs 42.6 ± 17.4, $t = -4.6$, $P < .001$). There was no difference between responders and nonresponders at baseline (39.2 ± 14.3 vs 42.1 ± 17.4, $t = -0.7$, not significant [NS]) The CUXOS scores of the responders significantly decreased from baseline to follow-up (paired $t = 4.3$, $P < .001$), whereas the scores of the nonresponders did not significantly change (paired $t = 0.6$, NS). At the follow-up visit there was a significant correlation between the CUXOS and the CGI-I rating ($r = 0.66$, $P < .001$) and the GAF ($r = -0.78$, $P < .001$).

A similar analysis was done for 43 patients with generalized anxiety disorder who completed the scale 5–22 weeks (mean = 11.7, SD = 3.0) after initiating treatment. Eighteen (41.9%) patients were rated much or very much improved on the CGI-I, and the mean GAF score was 58.9 (SD = 11.7). There was no difference between responders and nonresponders at baseline (35.9 ± 15.0 vs 35.3 ± 11.4, $t = 0.1$, NS), whereas patients who were much or very much improved scored significantly lower than the patients who were not treatment responders. (17.5 ± 13.9 vs 42.5 ± 13.1, $t = -6.0$, $P < .001$). The CUXOS scores of the responders significantly decreased from baseline to follow-up (paired $t = 3.8$, $P < .01$), whereas the scores of the nonresponders did not significantly change (paired $t = 1.0$, NS). At the follow-up visit there was a significant correlation between scores on the CUXOS and the CGI-I ($r = 0.61$, $P < .001$) and the GAF ($r = -0.74$, $P < .001$).

DISCUSSION

We believe that standardized scales should be routinely used to measure outcome when treating psychiatric disorders.[45] In fact, we believe that this should be the standard of care. Recently, the term *measurement-based care* has been coined in reference to the use of standardized scales to evaluate the outcome of treatment of depression.[46] If the standard of care is to change in the future, and scales are to be incorporated into clinical practice, then it will be necessary to consider feasibility issues as much as the psychometric properties of the measures.

The results of this large validation study of the CUXOS show that it is a reliable and valid measure of anxiety that is feasible to incorporate into routine clinical practice. On average, the scale takes approximately one and a half minutes to complete, and more than 90% of patients were able to complete it in less than 2 minutes. The CUXOS achieved high levels of internal consistency and test-retest reliability and was more highly correlated with other self-report measures of anxiety than with measures of depression, substance use problems, eating disorders, and anger, thereby supporting the convergent and discriminant validity of the scale. The CUXOS was also more highly correlated with blind interviewer ratings of the severity of anxiety than ratings of depression and anger. CUXOS scores were significantly different in patients with mild, moderate, and severe levels of anxiety, and patients with each of the *DSM-IV* anxiety disorders scored significantly higher than psychiatric patients without an anxiety disorder. Finally, the CUXOS was a valid measure of symptom change.

Our data and clinical experience allow us to approximate ranges of scores corresponding to a dimensional assessment of anxiety severity. We recommend that the nonanxious range corresponds to CUXOS scores of 0 to 10, minimal anxiety, 11–20; mild anxiety, 21–30; moderate anxiety, 31–44; and severe anxiety, 45 and above. Future studies should compare the CUXOS to a well-validated clinician measure of anxiety symptom severity such as the Hamilton Anxiety Rating Scale[4] to validate the score ranges.

There is no shortage of self-report questionnaires that assess anxiety; therefore, the development of any new scale should be questioned. The CUXOS distinguishes itself from existing instruments in several respects. Most anxiety scales measure the symptoms of a single anxiety disorder[23,25,47–52] or assess constructs underlying particular anxiety disorders.[18,24,33,53–55] In contrast, the CUXOS was intended as general measure of the severity of psychic and somatic anxiety. There are advantages and disadvantages toward this approach. Disorder-specific scales can only be used with patients with the index disorder, whereas a general anxiety measure can be useful for patients with any diagnosis who report symptoms of anxiety. Many depressed patients report high levels of anxiety in the absence of a diagnosable anxiety disorder.[27] Unpublished analyses of the MIDAS project data set likewise found that patients with substance use, adjustment, and somatoform disorders often received elevated scores on the SADS psychic and somatic anxiety items in the absence of a diagnosable anxiety disorder. Disorder-specific scales would not be appropriate in such situations.

Reliance on disorder-specific scales could be more time consuming and therefore more difficult to implement in clinical practice because patients with multiple anxiety disorders would need to complete multiple measures. Scale completion burden might interfere with the adoption of a measurement-based care approach toward treatment.

On the other hand, disorder specific scales may be more appropriate to characterize whether patients have remitted from a specific anxiety disorder. Future research should explore both clinicians' and patients' perspectives as to whether the use of general or disorder-specific scales is preferred. For phobic disorders, in particular, disorder-specific scales might be more valid because general scales assessing psychic and somatic symptoms of anxiety might underdetect ongoing pathology in patients who are able to avoid the phobic situations.

While there are other general measures of anxiety severity, some are either somewhat complicated to score[22] or expensive to purchase,[20] thereby making them less attractive to use in routine clinical practice. The CUXOS is the second in a series of clinically useful scales that we are developing for use in clinical practice. Previously, we described our development and validation of such a measure for depression, the CUDOS.[5,13,14] Each of the scales in the Clinically Useful series is intended to be brief, easily scored, and available to clinicians for personal use without cost. Each scale will have the same rating instructions thereby facilitating comparisons of symptom severity across varied symptom domains. While the CUXOS consists of both a psychic and a somatic anxiety factor, in our clinical practice, we compute only total scores because of the added time it takes to compute the factor scores. Perhaps if administered electronically it will be possible to simultaneously compute factor as well as total scores.

The CUXOS was designed to be brief and therefore more readily incorporated into routine clinical practice. We are not aware of any studies demonstrating that briefer scales are more likely to be used by clinicians than longer scales; however, a study[15] of depressed patients' acceptance of measurement-based care in clinical practice found that patients preferred to complete a briefer measure to monitor their progress.

In conclusion, the CUXOS is a reliable and valid brief self-administered anxiety questionnaire that can be incorporated into routine clinical practice without significant intrusion on patients,' clinicians,' or support staffs' time. While the results of this large validation study are encouraging, they require replication in samples with different demographic and clinical characteristics.

© Copyright 2010 Physicians Postgraduate Press, Inc.

Author affiliations: Rhode Island Hospital (Dr Zimmerman) and Department of Psychiatry and Human Behavior, Brown Medical School (all authors), Providence, Rhode Island.
Potential conflicts of interest: None reported.
Funding/support: None reported.

REFERENCES

1. Gilbody SM, House AO, Sheldon TA. Psychiatrists in the UK do not use outcomes measures: national survey. *Br J Psychiatry*. 2002;180(2):101–103.
2. Zimmerman M, McGlinchey JB. Why don't psychiatrists use scales to measure outcome when treating depressed patients? *J Clin Psychiatry*. 2008;69(12):1916–1919.
3. U.S. Department of Health and Human Services. Physician quality reporting initiative. In: *Center for Medicare and Medicaid Services*. 2008:110–275.
4. Hamilton M. The assessment of anxiety states by rating. *Br J Med Psychol*. 1959;32(1):50–55.
5. Zimmerman M, Chelminski I, McGlinchey JB, et al. A clinically useful depression outcome scale. *Compr Psychiatry*. 2008;49(2):131–140.
6. Zimmerman M. Integrating the assessment methods of researchers in routine clinical practice: The Rhode Island Methods to Improve Diagnostic Assessment and Services (MIDAS) project. In: First M, ed. *Standardized Evaluation in Clinical Practice*. Washington, DC: American Psychiatric Publishing, Inc; 2003:29–74.
7. Zimmerman M, Mattia JI. Psychiatric diagnosis in clinical practice: is comorbidity being missed? *Compr Psychiatry*. 1999;40(3):182–191.
8. Zimmerman M, Mattia JI. A self-report scale to help make psychiatric diagnoses: the Psychiatric Diagnostic Screening Questionnaire. *Arch Gen Psychiatry*. 2001;58(8):787–794.
9. Zimmerman M, Mattia JI. The reliability and validity of a screening Questionnaire for 13 DSM-IV Axis I disorders (the Psychiatric Diagnostic Screening Questionnaire) in psychiatric outpatients. *J Clin Psychiatry*. 1999;60(10):677–683.
10. Zimmerman M, Chelminski I. A scale to screen for DSM-IV Axis I disorders in psychiatric out-patients: performance of the Psychiatric Diagnostic Screening Questionnaire. *Psychol Med*. 2006;36(11):1601–1611.
11. Zimmerman M, Ruggero CJ, Chelminski I, et al. Developing brief scales for use in clinical practice: the reliability and validity of single-item self-report measures of depression symptom severity, psychosocial impairment due to depression, and quality of life. *J Clin Psychiatry*. 2006;67(10):1536–1541.
12. Zimmerman M, Sheeran T, Young D. The Diagnostic Inventory for Depression: a self-report scale to diagnose DSM-IV major depressive disorder. *J Clin Psychol*. 2004;60(1):87–110.
13. Zimmerman M, Posternak MA, Chelminski I. Using a self-report depression scale to identify remission in depressed outpatients. *Am J Psychiatry*. 2004;161(10):1911–1913.
14. Zimmerman M, Posternak MA, McGlinchey J, et al. Validity of a self-report depression symptom scale for identifying remission in depressed outpatients. *Compr Psychiatry*. 2006;47(3):185–188.
15. Zimmerman M, McGlinchey JB. Depressed patients' acceptability of the use of self-administered scales to measure outcome in clinical practice. *Ann Clin Psychiatry*. 2008;20(3):125–129.
16. Antony MM, Orsillo SM, Roemer L. *Practitioner's Guide to Empirically Based Measures of Anxiety*. New York, New York: Kluwer Academic/Plenum; 2001.
17. Burns GL, Keortge SG, Formea GM, et al. Revision of the Padua Inventory of obsessive compulsive disorder symptoms: distinctions between worry, obsessions, and compulsions. *Behav Res Ther*. 1996;34(2):163–173.
18. Chambless DL, Caputo GC, Jasin SE, et al. The Mobility Inventory for Agoraphobia. *Behav Res Ther*. 1985;23(1):35–44.
19. Turner SM, Beidel DC, Dancu CV, et al. An empirically derived inventory to measure social fears and anxiety: The Social Phobia and Anxiety Inventory. Psychological Assessment. *J Consult Clin Psychol*. 1989;1:35–40.
20. Beck AT, Epstein N, Brown G, et al. An inventory for measuring clinical anxiety: psychometric properties. *J Consult Clin Psychol*. 1988;56(6):893–897.
21. Foa EB. *Posttraumatic Stress Diagnostic Scale*. Minneapolis, MN: National Computer Systems; 1995.
22. Zung WW. A rating instrument for anxiety disorders. *Psychosomatics*. 1971;12(6):371–379.
23. Davidson JR, Miner CM, De Veaugh-Geiss J, et al. The Brief Social Phobia Scale: a psychometric evaluation. *Psychol Med*. 1997;27(1):161–166.
24. Meyer TJ, Miller ML, Metzger RL, et al. Development and validation of the Penn State Worry Questionnaire. *Behav Res Ther*. 1990;28(6):487–495.
25. Shear MK, Brown TA, Barlow DH, et al. Multicenter collaborative panic disorder severity scale. *Am J Psychiatry*. 1997;154(11):1571–1575.
26. Fawcett J. The detection and consequences of anxiety in clinical depression. *J Clin Psychiatry*. 1997;58(suppl 8):35–40.
27. Fava M, Rankin MA, Wright EC, et al. Anxiety disorders in major depression. *Compr Psychiatry*. 2000;41(2):97–102.
28. First MB, Spitzer RL, Williams JBW, et al. *Structured Clinical Interview for DSM-IV (SCID)*. Washington, DC: American Psychiatric Association; 1997.
29. Endicott J, Spitzer RL. A diagnostic interview: the schedule for affective disorders and schizophrenia. *Arch Gen Psychiatry*. 1978;35(7):837–844.
30. Zimmerman M, Mattia JI. Differences between clinical and research practices in diagnosing borderline personality disorder. *Am J Psychiatry*. 1999;156(10):1570–1574.
31. Garner DM, Olmstead MP, Polivy J. Development and validation of a multidimensional eating disorder inventory for anorexia nervosa and bulimia. *Int J Eat Disord*. 1983;2(2):15–34.
32. Beck AT, Rush AJ, Shaw BF, et al. *Cognitive Therapy of Depression*. New York, NY: The Guilford Press; 1979.
33. Leary MR. A brief version of the Fear of Negative Evaluation Scale. *Pers Soc Psychol Bull*. 1983;9(3):371–375.
34. Marks IM, Mathews AM. Brief standard self-rating for phobic patients. *Behav Res Ther*. 1979;17(3):263–267.
35. Foa EB, Riggs DS, Dancu CV, et al. Reliability and validity of a brief instrument for assessing post-traumatic stress disorder. *J Trauma Stress*. 1993;6(4):459–473.
36. Hodgson RJ, Rachman S. Obsessional-compulsive complaints. *Behav Res Ther*. 1977;15(5):389–395.
37. Selzer ML. The Michigan alcoholism screening test: the quest for a new diagnostic instrument. *Am J Psychiatry*. 1971;127(12):1653–1658.
38. Skinner HA. The drug abuse screening test. *Addict Behav*. 1982;7(4):363–371.
39. Pilowsky I. Dimensions of hypochondriasis. *Br J Psychiatry*. 1967;113(494):89–93.
40. Othmer E, DeSouza C. A screening test for somatization disorder (hysteria). *Am J Psychiatry*. 1985;142(10):1146–1149.
41. Swartz M, Hughes D, George L, et al. Developing a screening index for community studies of somatization disorder. *J Psychiatr Res*. 1986;20(4):335–343.
42. Guy W. *ECDEU Assessment Manual for Psychopharmacology*. Rockville, MD: National Institute of Mental Health; 1976.
43. Campbell DT, Fiske DW. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol Bull*. 1959;56(2):81–105.
44. Steiger JH. Tests for comparing elements of a correlation matrix. *Psychol Bull*. 1980;87(2):245–251.
45. Zimmerman M, McGlinchey JB, Chelminski I. An inadequate community standard of care: lack of measurement of outcome when treating depression in clinical practice. *Prim Psychiatry*. 2008;15:67–75.
46. Trivedi MH, Rush AJ, Wisniewski SR, et al. STAR*D Study Team. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. *Am J Psychiatry*. 2006;163(1):28–40.
47. Connor KM, Davidson JR, Churchill LE, et al. Psychometric properties of the Social Phobia Inventory (SPIN): new self-rating scale. *Br J Psychiatry*. 2000;176(4):379–386.
48. Houck PR, Spiegel DA, Shear MK, et al. Reliability of the self-report version of the panic disorder severity scale. *Depress Anxiety*. 2002;15(4):183–185.
49. Beidel DC, Borden JW, Turner SM, et al. The Social Phobia and Anxiety Inventory: concurrent validity with a clinic sample. *Behav Res Ther*. 1989;27(5):573–576.
50. Goodman WK, Price LH, Rasmussen SA, et al. The Yale-Brown Obsessive Compulsive Scale. I. Development, use, and reliability.

Arch Gen Psychiatry. 1989;46(11):1006–1011.
51. Blake DD, Weathers FW, Nagy LN. A clinician rating scale for assessing current and lifetime PTSD: The CAPS-1. *Behav Therapist*. 1990;13:187–188.
52. Foa E, Sachman L, Jaycox L, et al. The validation of a self-report measure of posttraumatic stress disorder: The Postraumatic Diagnostic Scale. *Psychol Assess*. 1997;9(4):445–451.
53. Chambless DL, Caputo GC, Bright P, et al. Assessment of fear in agoraphobics: the body sensations questionnaire and the agoraphobic cognitions questionnaire. *J Consult Clin Psychol*. 1984;52(6):1090–1097.
54. Harb GC, Heimberg RG, Fresco DM, et al. The psychometric properties of the Interpersonal Sensitivity Measure in social anxiety disorder. *Behav Res Ther*. 2002;40(8):961–979.
55. Peterson RA, Heilbronner RL. The Anxiety Sensitivity Index: construct validity and factor analytic structure. *J Anxiety Disord*. 1987;1(2):117–121.