A Scale for Rating Tricyclic Response in Major Depression: the TRIM

J. CRAIG NELSON, MD, AND CAROLYN M. MAZURE, PhD

Department of Psychiatry, Yale University School of Medicine, New Haven, Connecticut

We describe the construction and validation of the TRIM, an empirically derived scale designed to rate tricyclic antidepressant response in major depression. Symptoms selected were those that improved in direct association with therapeutic desipramine plasma levels, were frequently present, were substantially correlated with the scale total, could be reliably rated, and for which interview ratings were concordant with observed behavior. Eight symptoms met these criteria and were included. The sensitivity of the TRIM was tested and compared with the Hamilton Rating Scale for Depression (HAM-D) and the Montgomery Asberg Depression Scale (MADS) in a new 4-week prospective desipramine study of non-psychotic unipolar inpatients with major depression. TRIM 4-week totals were significantly associated with total desipramine plus hydroxydesipramine plasma concentrations, \( r = -0.32, p < 0.05 \), but HAM-D and MADS scores were not. Using multiple regression to control for pretreatment severity, TRIM scores were significantly associated with desipramine plus hydroxydesipramine levels, while HAM-D and MADS scores were not. The data appear to validate the sensitivity of the TRIM and illustrate that scales designed for drug response may detect drug effects that global scales do not.


THE HAMILTON RATING SCALE for depression (HAM-D),\(^1 \) its subsequent modifications,\(^2\text{-}^4 \) and other first-generation depression scales such as the Beck Depression Inventory,\(^5 \) Zung Depression Scale,\(^5 \) and Raskin Narcissistic Personality Inventory,\(^6 \) have been widely used in antidepressant drug studies. To some extent they have come to define or be equated with antidepressant response. Yet, these depression scales were designed to measure global change in the depressive state rather than drug effect per se. While these scales included symptoms that were thought to improve with drug treatment, it was not empirically established that each of the symptoms rated by these scales was specifically drug responsive.

More recently, other investigators have attempted to design scales that would be sensitive to or specific for antidepressant drug effect. Kelner\(^8 \) developed a brief depression scale for measuring change during drug treatment. His broadly defined scale items were based on inspection of prior drug-placebo studies but were not empirically derived. Montgomery and Asberg\(^9 \) reported the first empirically derived scale designed to measure antidepressant drug response (the Montgomery Asberg Depression Scale [MADS]). They selected scale items that correlated with global change in a heterogeneous sample treated with four different types of antidepressants. Although both scales were designed to measure drug response, the intent was to find common features of response to different antidepressant drugs.

Yet, there is increasing evidence that antidepressant drug effect is not a unitary construct. That is, a variety of drugs have antidepressant effects but differ in the types of patients or symptoms they treat. The tricyclic antidepressants (TCAs) remain the standard of treatment for severe or "endogenous" depression. The monoamine oxidase inhibitors (MAOIs), also effective antidepressants, appear to be more useful in depression with atypical features or prominent anxiety.\(^10\text{-}^12 \) Two decades ago, Overall and Hollister\(^13 \) questioned the specificity of the recognized antidepressants, noting that thioridazine and perphenazine were effective in some depressed patients, a finding recently reported for flupenthixol.\(^14 \) However, the neuroleptics do not appear to be as effective as the TCAs for symptoms central to depression, such as loss of interest.\(^15 \) The benzodiazepines have also been beneficial in anxious depressed outpatients but not as effective as TCAs in severe depression.\(^16 \) A related drug, alprazolam, is more effective in depressed outpatients than placebo\(^17 \) or diazepam,\(^18 \) but less effective than TCAs in endogenous depression.\(^19\text{-}^20 \) The introduction of several new drugs (e.g., trazodone, fluoxetine, nomifensine, and bupropion) that may have different profiles of...
action further complicates the definition of antidepressant drug action. In sum, the variety of drugs having some efficacy in depression suggests that the concept of a single unitary “antidepressant effect” may be too simplistic.

Rating scales designed to be more sensitive to specific pharmacologic profiles might help to define types of antidepressant effects. They might also help to detect more subtle drug effects, distinguish drug-related improvement in the presence of other symptoms, or determine true differences between drugs that global scales might fail to detect. The failure of global scales to detect true differences might in part explain the difficulty in demonstrating differences between TCAs and benzodiazepines or between TCAs and neuroleptics in depressed patients. Richter has noted that the HAM-D includes anxiety items that could result in a score of up to 18 but that are not unique to depression and do not differentiate antidepressant from anxiolytic effects. He suggested that the value of identifying “core depressive symptoms” to distinguish antidepressant drug effect.

**Development of a scale for TCA response in major depression (TRIM)**

We describe the development of a rating scale for the assessment of TCA response in major depression. We previously identified 10 symptoms that improved in direct association with therapeutic desipramine plasma levels. These symptoms were derived in patients having unipolar nonpsychotic major depression, since TCAs are effective in this syndrome and concurrent use of lithium or neuroleptics is not required. We also selected patients with a melancholic depression since this has been viewed as the “prototypic group” in which TCAs are employed and we wished to avoid inclusion of atypical patients. Relationships of blood levels and response are also better established in melancholic patients and rates of response to placebo, hospitalization, or other nonsomatic treatments are reduced in this group. Although our data were derived using one TCA, desipramine, the effects of the TCAs are reasonably similar when compared with the array of currently available antidepressant drugs.

In the current study we examined the 10 drug-responsive symptoms for inclusion in the scale. We required: (1) that symptoms were frequently present to assure that the scale reflected change in the severity of those symptoms, (2) that items significantly and substantially correlated with the scale total, (3) that symptoms could be reliably rated, and (4) that interview ratings could be validated by 24-hour observation. Using these guidelines, we constructed a scale for assessment of TCA response in major depression.

**Validation of the TRIM**

The TRIM was tested prospectively in a new sample of depressed patients to determine if the scale was more sensitive to TCA effect. The relationship of the TRIM to desipramine plasma levels and to an independent measure of global improvement was determined and compared with that found for the HAM-D and the MADS. We also examined the relationship of the TRIM with global severity and determined the internal consistency of the scale (i.e., its α-coefficient).

**Methods**

In section I, we described the development of the scale. In section II, we tested and validated the scale. Four samples were employed; sample I was used to derive the symptoms and determine their frequency and their correlation with the scale total. Sample II was used to assess the intrarater reliability of the items. The concordance of interview ratings with observed behavior was examined in sample III. Sample IV, a nonoverlapping sample, was used to validate the scale.

**Development of the scale**

**Determination of drug-responsive symptoms.** We previously determined the symptoms that significantly improved in relation to therapeutic desipramine plasma levels in 43 patients (sample I) having DSM-III melancholia who had not improved after a week of hospitalization, had a HAM-D score of at least 18, and were treated for a minimum of 3 weeks with desipramine 2.5 mg/kg. Individual symptom ratings on the Yale Depression Inventory were correlated with previously determined therapeutic desipramine plasma drug concentrations (above and below 115 ng/ml) using multiple regression analysis to control for pretreatment symptom severity. Ten symptoms that improved significantly (p < 0.05; Table 1) were considered for inclusion on the TRIM.

**Symptom frequency and severity.** The frequency and severity of the 10 drug-responsive symptoms was examined in sample I to assure that each item was frequently present and was rated across the range provided.

**Association of individual items with the scale total.** Pearson correlation coefficients were determined for each of the 10 symptoms and the scale total after 3 weeks of treatment in the 29 patients from sample I who had complete ratings on the Yale Depression Inventory. For this analysis, the symptom being correlated was deleted from the scale total. Stepwise multiple regression analysis was also used to determine the relative contribution of each of the 10 symptoms to the scale total.

**Interrater reliability.** The reliability of duplicate ratings was assessed using three raters who simultaneously rated 31 depressed inpatients with nonpsychotic major depression (sample II) using the Yale Depression Inventory. The interviewers performed the interviews in ro-
TABLE 1. Association of posttreatment symptom severity with therapeutic desipramine plasma concentrations in 43 patients.*

<table>
<thead>
<tr>
<th>Symptom</th>
<th>Standardized coefficient (b)</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Worthlessness</td>
<td>0.478</td>
<td>3.48</td>
<td>0.002</td>
</tr>
<tr>
<td>Decreased appetite</td>
<td>0.426</td>
<td>3.12</td>
<td>0.004</td>
</tr>
<tr>
<td>Depressed mood</td>
<td>0.425</td>
<td>3.17</td>
<td>0.003</td>
</tr>
<tr>
<td>Anhedonia†</td>
<td>0.387</td>
<td>2.31</td>
<td>0.03</td>
</tr>
<tr>
<td>Loss of interest and impaired performance</td>
<td>0.385</td>
<td>2.75</td>
<td>0.01</td>
</tr>
<tr>
<td>Hopelessness</td>
<td>0.366</td>
<td>2.56</td>
<td>0.01</td>
</tr>
<tr>
<td>Guilt</td>
<td>0.359</td>
<td>2.22</td>
<td>0.08</td>
</tr>
<tr>
<td>Scrotic anxiety</td>
<td>0.327</td>
<td>2.46</td>
<td>0.02</td>
</tr>
<tr>
<td>Decreased energy</td>
<td>0.320</td>
<td>2.35</td>
<td>0.02</td>
</tr>
<tr>
<td>Early morning awakening</td>
<td>0.255</td>
<td>2.02</td>
<td>0.05</td>
</tr>
<tr>
<td>Decreased concentration‡</td>
<td>0.297</td>
<td>1.75</td>
<td>0.09</td>
</tr>
<tr>
<td>Helplessness</td>
<td>0.249</td>
<td>1.63</td>
<td>0.11</td>
</tr>
<tr>
<td>Loss of sexual interest</td>
<td>0.240</td>
<td>1.62</td>
<td>0.11</td>
</tr>
<tr>
<td>Initial insomnia</td>
<td>0.237</td>
<td>1.65</td>
<td>0.11</td>
</tr>
<tr>
<td>Suicidal ideation</td>
<td>0.221</td>
<td>1.58</td>
<td>0.14</td>
</tr>
<tr>
<td>Middle insomnia</td>
<td>0.224</td>
<td>1.48</td>
<td>0.14</td>
</tr>
<tr>
<td>Psychotic anxiety</td>
<td>0.191</td>
<td>1.22</td>
<td>0.23</td>
</tr>
</tbody>
</table>

* Multiple regression used to account for pretreatment symptom severity. Separate regressions performed for each symptom.
† Symptoms assessed in 29 of the 43 patients.
‡ Symptom assessed in 40 of the 43 patients.

Notice that the intraclass correlations were determined for each of the raters.

Concordance of interview items with observed behavior. To determine if interview ratings were representative of the patient's behavior outside the interview, interview ratings of the 10 symptoms from the Yale Depression Inventory were compared with standardized nursing observations of ward behavior during a 3-day period concurrent with the interview in 31 depressed inpatients having nonpsychotic major depression (sample III). Some patients in sample III were included in sample II. Lack of concordance between the interview ratings and observed behavior would not necessarily mean the interview ratings were invalid since certain symptoms might be inherently subjective; however, a positive finding of concordance would support the validity of those items.

Evaluation of the scale

Association of the TRIM, the HAM-D, and the MADS with desipramine blood levels in a new patient sample. The relationship of the TRIM to plasma desipramine levels was examined in a new 4-week prospective study of depressed inpatients (sample IV) and compared with that for the HAM-D and the MADS. Consecutive inpatients who had DSM-III unipolar nonpsychotic major depression and who remained depressed after 1 week of hospitalization without antidepressant treatment began a 4-week desipramine trial. Patients with schizophrenia, orga-
Table 2. Frequency, severity, and correlation of individual symptoms with the scale total in patients in sample I

<table>
<thead>
<tr>
<th>Symptom</th>
<th>Frequency (N = 43)</th>
<th>Severity (N = 48)</th>
<th>Correlation with the scale total (N = 29)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%</td>
<td>Mean</td>
<td>Range</td>
</tr>
<tr>
<td>Depressed mood</td>
<td>100</td>
<td>2.59</td>
<td>1-4</td>
</tr>
<tr>
<td>Hopelessness</td>
<td>95</td>
<td>2.07</td>
<td>0-4</td>
</tr>
<tr>
<td>Anhedonia</td>
<td>95*</td>
<td>2.57</td>
<td>0-4</td>
</tr>
<tr>
<td>Worthlessness</td>
<td>91</td>
<td>1.82</td>
<td>0-4</td>
</tr>
<tr>
<td>Loss of interest and impaired performance</td>
<td>100</td>
<td>2.71</td>
<td>1-4</td>
</tr>
<tr>
<td>Guilt</td>
<td>80</td>
<td>1.84</td>
<td>0-3</td>
</tr>
<tr>
<td>Decreased appetite</td>
<td>78</td>
<td>1.14</td>
<td>0-2*</td>
</tr>
<tr>
<td>Somatic anxiety</td>
<td>87</td>
<td>1.42</td>
<td>0-4</td>
</tr>
<tr>
<td>Decreased energy</td>
<td>93</td>
<td>1.60</td>
<td>0-2*</td>
</tr>
<tr>
<td>Early morning awakening</td>
<td>77</td>
<td>1.44</td>
<td>0-2*</td>
</tr>
</tbody>
</table>

*p < 0.05, **p < 0.01, ***p < 0.001.
* The 29 patients with ratings for all symptoms including anhedonia.
+ Rated in 29 of the 43 patients.

Sixty-five patients with nonpsychotic unipolar major depression began desipramine treatment. Of these, nine patients developed side effects that interrupted treatment, seven patients left the hospital prematurely for nonmedical reasons, three patients worsened and treatment was stopped, and OH-DMI levels could not be determined in three patients due to assay interference. Thus, 43 patients (28 women and 15 men) completed the trial and were included in the analysis. Their mean age was 48 ± 16 years, the mean duration of the episode was 39 ± 78 weeks, and the mean initial HAM-D was 24 ± 4.8.

Association of the TRIM and HAM-D to global severity. The relationship of the TRIM to global severity on the CGI was compared with that for the HAM-D and the MADS in sample IV using the ratings before and after drug treatment. The linearity of the association of TRIM scores and global severity—a characteristic described by Bech and coauthors—as "ascending monotonicity"—was also assessed by plotting mean TRIM scores at each level of global severity encountered after 1 week of hospitalization without antidepressants. The sample of 88 included the 65 patients in sample IV who began desipramine treatment as well as patients with major depression who responded after 1 week.

Internal consistency. The α-coefficient, a measure of internal consistency, was determined and compared with that of the HAM-D in sample IV using 4-week ratings.

Results

Development of the scale

Symptom frequency and severity. All 10 drug-responsive symptoms were present in at least 73% of the patients (Table 2). These symptoms varied across the full range of severity except for guilt, which ranged from 0-3; however, a rating of 4 required the patient to be delusional.

Association of individual items with the scale total. For five symptoms—hopelessness, loss of interest/impaired performance, anhedonia, worthlessness, and depressed mood—the magnitude of the correlations with the scale total was substantial (r > 0.6, p < 0.001; Table 2). The correlations of the scale total with guilt and decreased appetite were moderately strong (r = 0.56, p < 0.01). Somatic anxiety and loss of energy were less strongly correlated with the total (r = 0.44 and 0.42, p < 0.05). Early-morning awakening was not significantly associated with the scale total.

When the drug-responsive symptoms were regressed on the scale total, hopelessness, loss of interest in work and activities, worthlessness, decreased appetite, depressed mood, decreased ability to experience pleasure, and guilt entered stepwise and accounted for 97.5% of the variance. Somatic anxiety, early-morning awakening, and loss of energy added less than 3% to the variance in the scale total once the seven other symptoms entered the regression.

Interrater reliability. The reliability of the drug-responsive symptoms (Table 3) appeared to be adequate with intraclass correlations of 0.70 or better. Loss of interest and impaired performance were assessed separately rather than as the combined HAM-D item for loss of interest in work and activities.

Concordance of interview ratings with ward behavior. The concordance of interview ratings of the drug-responsive symptoms with observation of the behavior is shown in Table 3. Loss of interest and impaired performance were assessed separately. Sleep difficulty, somatic anxiety, and loss of energy were not concordant.

Construction of the scale. Three of the 10 drug-responsive symptoms—early-morning awakening, somatic
Table 3. Interrater reliability and concordance with observed behavior

<table>
<thead>
<tr>
<th>Symptom</th>
<th>Interrater reliability, intraclass correlations (N = 31)</th>
<th>Concordance with observed behavior (N = 31)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depressed mood</td>
<td>0.70***</td>
<td>0.58***</td>
</tr>
<tr>
<td>Hopelessness</td>
<td>0.70***</td>
<td>0.58***</td>
</tr>
<tr>
<td>Anhedonia</td>
<td>0.74***</td>
<td>0.61**</td>
</tr>
<tr>
<td>Worthlessness</td>
<td>0.80***</td>
<td>0.40**</td>
</tr>
<tr>
<td>Loss of interest</td>
<td>0.85***</td>
<td>0.49**</td>
</tr>
<tr>
<td>Impaired performance</td>
<td>0.81***</td>
<td>0.49**</td>
</tr>
<tr>
<td>Guilt</td>
<td>0.85***</td>
<td>0.48**</td>
</tr>
<tr>
<td>Decreased appetite</td>
<td>0.94***</td>
<td>0.65***</td>
</tr>
<tr>
<td>Somatic anxiety</td>
<td>0.95**</td>
<td>0.28</td>
</tr>
<tr>
<td>Decreased energy</td>
<td>0.85***</td>
<td>0.27</td>
</tr>
<tr>
<td>Early-morning awakening</td>
<td>0.80***</td>
<td>0.34</td>
</tr>
</tbody>
</table>

*p < 0.05; **p < 0.01; ***p < 0.001.

anxiety, and loss of energy — were weakly correlated with the scale total and were not concordant with observed behavior. These symptoms were also weakly associated with drug levels. Exclusion of these three items had little effect on the performance of the remaining items of the scale, which accounted for 97% of the variance in the scale total. Further, there was no difference in the correlation of the scale with global severity whether or not these three items were included. For these reasons, these three symptoms were dropped from the scale.

The HAM-D item "decreased interest in work and activities" was highly correlated with drug levels, the scale total, and global severity. Yet, this item rates two different concepts — decreased interest and impaired performance. We found patients commonly described that they "pushed themselves" to perform a task even if they had little interest in it. Psychological tests of impaired correlation confirmed with patients' reports of impaired task performance (unpublished data), giving support to the validity of the impaired performance item. Loss of interest and impaired performance correlated substantially with the scale total, were present in at least 87% of the cases, could be reliably rated, and were concordant with observation. Thus, we included decreased interest and impaired performance as separate ratings on the scale.

To give items equal weight, each was rated on a 0- to 4-point scale (including decreased appetite), with appropriate anchor points devised. With the modifications described, an eight-item scale was constructed (Appendix).

Evaluation of the scale

Association of the TRIM, the HAM-D, and the MADS with desipramine blood levels in a new patient sample.

Table 4. Correlations of total desipramine + OH-DMI concentrations with response on the TRIM, HAM-D, and MADS

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>r*</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRIM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Response/nonresponse</td>
<td>43</td>
<td>0.22</td>
<td>0.07</td>
</tr>
<tr>
<td>4-week TRIM</td>
<td>36</td>
<td>-0.22</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>HAM-D</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Response/nonresponse</td>
<td>41</td>
<td>0.02</td>
<td>NS</td>
</tr>
<tr>
<td>4-week HAM-D</td>
<td>36</td>
<td>-0.20</td>
<td>NS</td>
</tr>
<tr>
<td>MADS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Response/nonresponse</td>
<td>41</td>
<td>0.09</td>
<td>NS</td>
</tr>
<tr>
<td>4-week MADS</td>
<td>36</td>
<td>-0.28</td>
<td>&lt;0.10</td>
</tr>
</tbody>
</table>

* Point biserial correlations for response/nonresponse and total drug levels. Spearman rank order correlations for week 4 scale totals and drug levels.

The correlations of total drug concentrations (desipramine plus OH-DMI) with response/nonresponse and 4-week scores on the TRIM, the HAM-D, and the MADS, were determined in 43 patients (sample IV, Table 4). Total desipramine plus OH-DMI concentrations were significantly correlated with 4-week TRIM scores and marginally correlated with the TRIM cutoff (r = 0.23, p = 0.07). Correlations of total drug concentrations and HAM-D measures were low (0.02 and -0.20) and not significant. The correlation of the MADS cutoff with drug levels was low (r = 0.09), but that for the 4-week MADS scores approached significance (r = -0.28, p < 0.10).

The association of 4-week scores and total drug levels was also examined using multiple regression to account for the pretreatment scale severity (Table 5). Four-week TRIM scores were significantly associated with total drug levels when the pretreatment TRIM score was accounted for, but the 4-week HAM-D and MADS scores were not related to drug levels using this method.

Association of the scales with global improvement. Correlations of global improvement with 4-week total scores were examined and were -0.87 for the TRIM, -0.76 for the HAM-D, and -0.86 for the MADS (p < 0.001 for each, N = 36).

Association with global severity. After 1 week of hospitalization but before drug treatment, the correlations of global severity with TRIM, HAM-D, and MADS totals were 0.60, 0.79, and 0.83, respectively (p < 0.001 for each, N = 89). Following treatment the correlations of the TRIM, HAM-D, and MADS with global severity were 0.82, 0.85, and 0.89 (p < 0.001 for each, N = 36). As depicted in Figure 1, mean predrug TRIM scores rose in linear fashion in relation to global severity.

Internal consistency. The coefficient for the eight symptoms of the TRIM in sample IV patients who completed the trial was 0.85. For comparison, the coefficient for the HAM-D in the same group was 0.84.
A Scale for Tricyclic Response

Table 5. Association of 4-week scale scores and desipramine + OH-DMI levels using multiple regression to account for pretreatment severity.

<table>
<thead>
<tr>
<th>Pretreatment score</th>
<th>Total drug level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \hat{\beta} )</td>
</tr>
<tr>
<td></td>
<td>( \hat{\beta} )</td>
</tr>
<tr>
<td>TRIM</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>-0.32</td>
</tr>
<tr>
<td>HAM-D</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>-0.10</td>
</tr>
<tr>
<td>MADS</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>-0.21</td>
</tr>
</tbody>
</table>

* Separate regressions were performed for each scale. Pretreatment scale scores and total drug concentrations were regressed on the 4-week scale scores.

Discussion

We have developed an eight-item scale (the TRIM) to assess symptom improvement, which is directly related to TCA effect in major depression. In a prospective validation study, the TRIM was significantly associated with total desipramine plus OH-DMI plasma levels, while the HAM-D and the MADS were not. This finding supports the validity of the TRIM and illustrates how the use of more global scales (such as the HAM-D) might fail to detect a plasma level-response relationship that the TRIM identified.

Although the magnitude of the correlations for drug levels and response were not substantial, this may reflect the narrow distribution of plasma levels using the dose-adjustment method. The correlation of response on the TRIM with drug levels increased and became significant if a desipramine plus OH-DMI threshold of 217 ng/ml was used (\( \rho = 0.34, p = 0.01 \)). However, this threshold was based on inspection of the data, while total concentration expressed as a continuous variable was unbiased. HAM-D measures were not significantly associated with drug levels even if this threshold was employed.

The differences between the correlations of blood levels and response on the three scales did not differ significantly, although the comparison of the correlations for plasma levels and response on the TRIM and HAM-D approached significance. However, as Montgomery and Asberg\(^{44}\) noted, differences of this magnitude may be sufficient to demonstrate a significant relationship using one scale but not the other, which was the case in this study.

HAM-D items excluded from the scale. Certain HAM-D items not included in our scale—hypochondriasis, lack of insight, agitation, loss of weight, suicidal ideation, and decreased sexual interest—were either infrequent or not closely associated with change during drug treatment, as others have noted.\(^{8, 35}\) In fact, inclusion of weight loss and loss of sexual interest can be misleading in hospitalized patients. On admission, both are scored using prior history and are frequently present. However, in the hospital, weight loss results in intervention and is rare; and some patients report that loss of sexual interest is difficult to assess in the hospital. As a result, during treatment, both symptoms may be rated absent without a meaningful change in the patient.

Loss of energy and somatic anxiety were modestly drug responsive, but added little to the scale. Motor retardation, a cardinal symptom of endogenous depression,\(^{37}\) is useful for diagnosis if present, but was not useful for measuring change since it was not frequent in these nonpsychotic unipolar patients—a finding noted by others.\(^{9}\)

Although reports of sleep disturbance are common in depression, and prior studies\(^{38, 39}\) suggest that sleep is the first symptom to respond to amitriptyline, the effects on sleep of other effective antidepressants—desipramine,\(^{40}\) tranylcypromine, phenelzine, and fluoxetine\(^{41}\)—are variable. If effective antidepressant drugs differ substantially in their effects on sleep, it seems questionable whether their hypnotic effects are central to antidepressant action. Other factors confounding the rating of sleep include: (1) patients' reports of sleep disturbance may be inaccurate, (2) sleep disturbance may occur as a symptom of anxiety disorder, and (3) depressed patients commonly receive hypnotic drugs that confound the use of this item. We conclude that there are substantial reasons for separating the assessment of sleep disturbance from scales specific for antidepressant drug effect.

Comparison with other specific scales. Bech and coauthors\(^{42}\) have described an endogenous subscale of the HAM-D intended to measure change. The six items of the subscale were selected on the basis of their frequency and their linearity and dispersion in relation to global severity; however, a direct association of these items with drug treatment was not examined. One difference in their scale was inclusion of motor retardation, which we found to be infrequent. An endogenous subscale of the HAM-D, sometimes referred to as a “core depression” scale, has also been described,\(^{42, 43}\) but it was de-

![Fig. 1. Mean TRIM totals for patients with different levels of global severity. Mean TRIM scores appear to rise in linear association with global severity.](image-url)
signed for diagnosis of endogenicity, not for measuring change.

Kellner's Brief Depression Rating Scale was developed to assess change during drug treatment. The scale was designed based on a review of drug-placebo and low-dose-high-dose studies, but was not empirically derived or specific for one class of antidepressants. In contrast to the TRIM and the MADS, the eight items of Kellner's scale were broadly defined.

The first empirically derived scale intended for the assessment of antidepressant response (the MADS) was developed by Montgomery and Asberg. They selected 10 items that occurred frequently and were highly correlated with global change during drug treatment. The primary difference between their scale and ours is that their intent was to find common features of drug response in heterogeneous patients receiving different classes of drugs; ours was to define characteristics of TCA response in typical endogenous patients.

Perhaps the most important difference between our scale and the MADS is the inclusion of psychic anxiety on the MADS. We found psychic anxiety to have one of the lowest correlations with therapeutic drug levels of any of the symptoms studied (Table 1). Consistent with this finding, high pretreatment anxiety has been reported to predict poor response to imipramine and amitriptyline.

Anxiety may also be crucial for distinguishing TCA effect from other drugs. Davidson and colleagues recently reported that psychic anxiety was one of the most responsive symptoms to the MAOI isocarboxazid, a finding consistent with prior reports. Psychic anxiety may also respond to benzodiazepines and neuroleptics. Exclusion of psychic anxiety from the TRIM would appear to be particularly important for distinguishing TCA response from other drug profiles.

Generalizability of the scale. It is reasonable to question whether our scale, developed with desipramine, will apply to the other TCAs. The items included on the TRIM primarily reflect mood, interest, or ideation and would appear to be responsive to all the TCAs and not likely to be affected by secondary drug effects such as sedation. The one somatic symptom that may respond differently to other TCAs is loss of appetite. Change in this symptom is complicated by the interaction of change in the depression and direct drug effects on appetite. Desipramine is an intermediate drug with respect to this symptom; it has little effect on appetite itself but is associated with weight gain in patients whose depression improves. Other TCAs would be expected to have a greater effect on appetite.

Although we selected melancholic patients to derive scale items in order to assure that we had a "prototypic" group of TCA-responsive patients, our purpose was to design a scale for assessment of TCA response in major depression. Our data indicate that the TRIM is more sensitive than the HAM-D for detecting change related to desipramine treatment in patients with major depression. Nevertheless, major depression is a heterogeneous diagnostic group and it remains unclear how to best delineate those patients who are TCA responsive. There is evidence that mood-reactive atypical depression, as defined by the Columbia University group, is less responsive to TCAs and might be excluded from this group.

Whether this scale is more specific for TCA effects will require comparison studies with other drugs. If the TRIM has a more narrow focus than global scales, we would anticipate that it would not replace global scales but would be used with them and would complement them.

Applications of the scale. A scale specific for TCA response might have a variety of applications. It might help to establish if new drugs have a profile similar to that of the TCAs. The TRIM may provide a sharper focus for comparisons between the effects of the TCAs, MAOIs, anxiolytic drugs, or neuroleptics in depression. It may be useful in mixed psychiatric syndromes, e.g., double depression, to separate drug-responsive symptoms from other features of the syndrome that might obscure drug response on global scales. This scale might also be useful in the assessment of medically ill depressed patients in whom somatic symptoms are less useful. Cognitive or psychological symptoms may be more helpful in assessing depression in these patients and the TRIM appears to measure these symptoms.

From a conceptual perspective, response to antidepressant medications remains one of our best means of exploring the "pharmacologic bridge" between the neurochemical mechanisms of drug action and the symptoms that drugs treat. While the development of more selective and specific antidepressants should facilitate this exploration, scales that provide sharper behavioral definition may be necessary to measure their effects.

Acknowledgments

This study was supported in part by National Institute of Mental Health grant 1 P50 MH30020 through the Yale Mania/Health Clinical Research Center, New Haven, Connecticut, and by a grant from Merrell Dow Pharmaceuticals.

References

4. Carroll RJ, Feinberg M, Smouse PE, Rawson SG, Greidanus J. The
A Scale for Tricyclic Response


APPENDIX

A Scale for Tricyclic Response in Major Depression: The TRIM

1. Depressed mood

Q: How have your spirits been during the past week? Have you felt blue, down in the dumps, depressed? How would you describe it? How bad has it been? Have you wanted to cry? Or felt so badly that you couldn’t cry? (If the patient cannot respond to these questions adequately, ask: What has your mood been like?)

I: Depressed mood includes reported feelings and appear-
2. Appetite

Q: Have you noticed a decrease in your appetite? Are you still eating a normal amount?
I: Only rate decrease in appetite
0 = No change
1 = Doubtful, trivial
2 = Mild, decreased interest in food, normal intake
3 = Moderate, little interest in food, decreased intake
4 = Severe, no interest in food, does not initiate eating

3. Loss of interest

Q: Do you look forward to work or activities, or do you find you have lost interest in these things?
I: Rate the anticipation of pleasure in activities or work. It may be helpful to elicit what activities patient usually enjoys.
0 = Usual interest
1 = Slight or doubtful decrease in interest
2 = Mild, less interest, less apt to initiate activity
3 = Moderate, little interest in work or activity
4 = Severe, no interest in any pleasurable activity

4. Loss of ability to experience pleasure

Q: Are you able to enjoy things the way you usually do? For example, during the past week if you visited someone, watched TV, or read a book . . . (other) . . . did you enjoy it?
I: Again, as in 3, it may be necessary to elicit from the patient what is usually enjoyable, i.e., someone visiting, seeing the children, reading, watching TV. Include response to efforts of others to engage the patient.
0 = Usual ability to experience pleasure
1 = Slight or doubtful decrease in ability to experience pleasure
2 = Mild, definite, or spontaneously reported decrease
3 = Moderate, little pleasure even in favorite things
4 = Severe, no pleasure in anything

5. Impaired performance

Q: Has it been more difficult to do your work, household chores, or other tasks? Do you have to push yourself to finish tasks?
I: Rate performance of work, household tasks, or activities of daily living. Include personal hygiene here.
0 = No difficulty
1 = Slight or doubtful
2 = Mild, definite decrease in work or chores performed
3 = Moderate, cannot work outside of home; functioning limited to a few household chores; in hospital performs only required tasks
4 = Severe, in or out of hospital, completes no tasks, personal hygiene neglected, spends much time in bed

6. Pathological guilt

Q: Have you been feeling guilty about events in your life or something you’ve done? Have you been thinking you're a bad or evil person?
I: Distinguish from feelings of failure. Guilty persons should feel there is an element of badness, evil, or sinfulness in themselves or their actions.
0 = Absent
1 = Doubtful, only mentioned on questioning, expresses self-reproach but does not state he/she is "bad"
2 = Ideas stated spontaneously or are clearly present; patient thinks he/she has “sinned,” wronged others, or is bad
3 = Ideas prominent, illness may be punishment
4 = Delusional guilt, i.e., has harmed others, serious consequences of guilt imminent; hears accusatory voices

7. Worthlessness/failure

Q: Have you been down on yourself? Do you feel you’ve failed?
I: Distinguish from guilt, but may be present with guilt.
0 = Absent
1 = Doubtful
2 = States definite feelings of failure
3 = Prominent concern with failure or spontaneously stated feelings of failure
4 = Delusions of worthlessness

8. Hopelessness

Q: What kind of future do you see for yourself? Do you think things will get better?
0 = Hopeful things will get better
1 = Doubtful about future but not clearly hopeless
2 = Volunteers hopeless feelings, feels quite hopeless at times
3 = Persistent hopelessness, cannot be dispelled
4 = Nihilistic delusions, i.e., patient is doomed, world is ending

* Q. question or probe; I, rating instruction.