Predictors of Response and Nonresponse to Light Treatment for Winter Depression

Michael Terman, Ph.D., Leora Amira, Ph.D.,
Jiuan Su Terman, Ph.D., and Donald C. Ross, Ph.D.

Objective: The authors' goal was to determine whether the pattern and severity of depressive symptoms predict response to light treatment for seasonal affective disorder. Method: Subjects with winter depression (N=103) were given bright light treatment. Seventy-one were classified as responders, 15 as nonresponders, and 17 as partial responders. Using depression rating scale data and correlational and multivariate analysis, the authors sought predictors of response in baseline symptom and scale scores. Results: Responders were characterized by atypical symptoms, especially hypersomnia, afternoon or evening slump, reverse diurnal variation (evenings worse), and carbohydrate craving. By contrast, nonresponders were characterized mainly by melancholic symptoms: retardation, suicidality, depersonalization, typical diurnal variation (mornings worse), anxiety, early and late insomnia, appetite loss, and guilt. The ratio of atypical to classical symptoms of depression, rather than severity per se, best predicted treatment outcome for the group as a whole. Pretreatment expectations were positively correlated with improvement on the Hamilton Depression Rating Scale but not on a supplementary scale of atypical symptoms. Conclusions: Light-responsive seasonal affective disorder is distinguished by a dominant atypical symptom profile closely associated with depressed mood. Nonresponders form a clinically distinct group with melancholic features. The patient's symptom profile, therefore, should be considered when diagnosing seasonal affective disorder and selecting treatment.

Since the original clinical trial of artificial bright light for treatment of winter depression in seasonal affective disorder (1), the method has seen widespread use. In a cross-center analysis (2), results for more than 300 patients showed remission rates of about 50% for light therapy in the morning or in the morning and the evening. A substantial number of patients failed to respond or showed only partial improvement. The question remains as to whether responders and nonresponders can be distinguished by their baseline symptom profiles or other clinical features. Predictors would serve not only to refine the definition of the disorder (3) (which was identified in DSM-III-R solely on the basis of a recurrent seasonal depressive pattern) but also guide clinicians toward alternate treatments for likely nonresponders and aid research on the mechanism of action of light.

Several studies have used the Structured Interview Guide for the Hamilton Depression Rating Scale—Seasonal Affective Disorder Version (4) to identify symptoms that correlate with treatment response. This instrument merges the Hamilton Depression Rating Scale (5) with a supplementary 8-item scale for atypical symptoms, including prominent symptoms of hypersomnia, hyperphagia, carbohydrate craving, and afternoon or evening slump, Stinson and Thompson (6), for example, found a positive correlation between baseline scores on the atypical symptom scale and percent of improvement but negative correlations between baseline Hamilton scale scores and improvement on the atypical scale. Similarly, Nagayama et al. (7) found a positive correlation between baseline scores on the atypical scale but not the Hamilton scale and improvement on the two scales combined.

Others have identified specific symptoms as potential predictors of response to light. Oren et al. (8) found that hypersomnia, carbohydrate craving, and suicidality were positive predictors in a multiple regression analysis using change scores on the Hamilton Depression Rating Scale—Seasonal Affective Disorder Version. Lam (9) found such correlations for hypersomnia, in-
increased eating, and younger age. Using prospective ratings, Kräuchi et al. (10) found that consumption of sweets in the afternoon and evening, but not diurnal mood variation or sleep-related symptoms, predicted response. Similarly, Meesters et al. (11) found no predictive value in diurnal variation or hypersomnia. Although these studies offer suggestive leads, conclusions vary. Underlying problems include small sample sizes, inconsistent criteria for subject inclusion and measures of posttreatment change, contrasting statistical analyses, and lack of relapse during withdrawal (12).

Previous predictor studies have emphasized multiple regression analysis of symptom ratings, which assumes a homogeneous group of subjects. However, recurrent winter mood disorders may comprise distinct clinical subgroups that are not discriminated by current criteria for seasonal affective disorder (13). Multivariate analyses offer an alternate approach. In dichotomous analyses, one can search for features that best discriminate responders from nonresponders. In intercorrelational analyses (such as cluster analysis, factor analysis, and multidimensional scaling), one can explore patterns of distinctive features or traits; the novelty of associated traits in a multidimensional distribution points to sources of subject heterogeneity. The present study compares correlational and multivariate approaches with the aim of a convergent validation of predictive factors.

METHOD

Subjects

Subjects were 103 research volunteers enrolled between 1985 and 1992, including 83 women and 20 men who ranged in age from 18 to 63 years (mean=39.0, SD=10.2). Diagnosed mood disorders were major depression, recurrent (DSM-III-R 296.3) (N=74), bipolar disorder, depressed (DSM-III-R 296.5) (N=2), bipolar disorder not otherwise specified (DSM-III-R 296.7) (N=16), and depressive disorder not otherwise specified (DSM-III-R 311.0) (N=11), all with seasonal pattern (winter type). Subjects also met criteria of the National Institute of Mental Health (NIMH) for seasonal affective disorder (1), except for the group with DSM-III-R diagnosis of depressive disorder not otherwise specified, who, however, met equivalent severity criteria for current episode. Results of physical examinations, complete blood counts, urinalysis, and ECG for all subjects were normal. Subjects were without cataracts, glaucoma, or retinal disease, were free of antidepressant medications, and had no previous experience with light treatment.

Procedures

Assessment instruments included the Hamilton Depression Rating Scale—Seasonal Affective Disorder Version (4) (the combined scale) and the Clinical Global Impression (CGI) (14); evaluations were made by raters blind to treatment status. Subjects entered treatment after providing written informed consent. The baseline score on the combined scale of all subjects was at least 20. Their score on the Hamilton scale alone was at least 10 and their score on the supplementary scale for atypical symptoms was at least 5 (15). In the last 3 years of the study, 59 of the subjects also provided expectation ratings on a 7-point scale (1=low, 4=neutral, 7=high). For the first 2 years of the study, a 2500-lux fluorescent light fixture was used. The subjects received 2-hour daily home-treatment sessions for 7 to 10 days in the morning, in the evening, or both. After 2 years, subjects received crossovers using 10,000 lux (16) and 30-minute sessions in the morning or evening for 10–14 days. Compliance was monitored by telephone log-in.

Responders were identified as subjects whose posttreatment score on the combined scale was reduced by at least 50% (with both Hamilton and atypical scores of 7 or lower) and CGI ratings of 1 (very much improved) or 2 (much improved). Seventy-one (68.9%) of the 103 subjects were classified as responders. When crossovers were given, the best response on either side of the crossover was used (16). Data were included only if relapse occurred within 3 weeks of withdrawal. Nonresponders were identified as subjects whose posttreatment scores on the combined scale increased, or decreased less than 25%, and whose CGI ratings were 4 (unchanged) or higher (worsened). Fifteen (14.6%) of the subjects were classified as nonresponders. This classification was further verified by lack of response to longer light exposure sessions (45–60 minutes) during follow-up. Partial responders were identified as subjects with score reductions on the combined scale between 30% and 50% and CGI ratings of 3 (minimal improvement). Seventeen (16.5%) of the subjects were classified as partial responders.

Data Analysis

Atypical balance. Beyond the symptom severity measures of the Hamilton and atypical scales, the relative dominance of atypical symptoms was expressed as a percentage (atypical score combined score)×100.

Effect sizes (17). Between-group comparisons of baseline symptom severity for items on the combined scale were expressed as the standardized difference between means (d): d=0.3 was considered small; d=0.5, moderate; and d=0.8, large. The effect size for the relation of pretreatment expectaions ratings to posttreatment response was evaluated by the Pearson correlation coefficient (r): r=0.1, small; r=0.3, moderate; and r=0.5, large. Between-group comparisons of historical factors (e.g., use of medication) were evaluated by the effect size of proportions (h): h=0.3, small; h=0.5, moderate; h=0.8, large.

Discriminant analysis (18). Based on the pretreatment symptom scores of responders and nonresponders, we used a forward stepwise discriminant analysis (F to enter=0.0, F to remove=3.996) to derive an equation describing the canonical variable that best distinguished the groups, followed by a jackknifed classification to determine the proportion of cases accounted for.

Intercorrelational analyses (19). A hierarchical cluster analysis based on Euclidean distances and the average linkage method was used to compare the intercorrelation of baseline symptoms in responders and nonresponders separately. The resulting tree diagrams were compared with multidimensional scaling and principal component factor analyses.

Stepwise linear regression (20). A stepwise linear regression using percent change scores on the combined scale, and individual symptom and scale scores, was performed. The value of F to enter or remove was between 4.0 and 3.996, respectively, and p<0.05 was required to enter and remain in the regression.

RESULTS

Contrasts Between Responders and Nonresponders at Baseline

Rating scale summary. Although scores on the combined scale did not differ between the groups at baseline, responders showed a significantly lower mean Hamilton score (mean=15.51, SD=3.64, versus mean=20.20, SD=5.13) (t=-4.20, df=84, p<0.001, two-tailed), and their atypical score was slightly higher (mean=14.23, SD=4.35, versus mean=12.00, SD=5.58) (t=-1.71, df=84, p<0.09). This resulted in significantly higher
TABLE 1. Baseline Frequency of Depressive Symptoms Among Subjects With Winter Depression Who Did or Did Not Respond to Light Treatment

<table>
<thead>
<tr>
<th>Measure From Hamilton Depression Rating Scale or Supplementary Scale for Atypical Symptoms</th>
<th>Effect Size (d)</th>
<th>Responders (N=71)</th>
<th>Nonresponders (N=15)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atypical balance score</td>
<td>1.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypersomnia (scale for atypical symptoms)</td>
<td>0.79</td>
<td>57 80</td>
<td>6 40</td>
</tr>
<tr>
<td>Afternoon or evening slump (scale for atypical symptoms)</td>
<td>0.63</td>
<td>65 91</td>
<td>12 80</td>
</tr>
<tr>
<td>Evenings worse (scale for atypical symptoms)</td>
<td>0.51</td>
<td>27 38</td>
<td>2 13</td>
</tr>
<tr>
<td>Total score on atypical symptoms</td>
<td>0.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carbohydrate craving (scale for atypical symptoms)</td>
<td>0.46</td>
<td>62 87</td>
<td>11 73</td>
</tr>
<tr>
<td>Increased eating (scale for atypical symptoms)</td>
<td>0.39</td>
<td>51 72</td>
<td>8 53</td>
</tr>
<tr>
<td>Weight gain (scale for atypical symptoms)</td>
<td>0.24</td>
<td>42 59</td>
<td>9 60</td>
</tr>
<tr>
<td>Increased appetite (scale for atypical symptoms)</td>
<td>0.17</td>
<td>51 72</td>
<td>10 67</td>
</tr>
<tr>
<td>Fatigability (scale for atypical symptoms)</td>
<td>0.03</td>
<td>69 97</td>
<td>15 100</td>
</tr>
<tr>
<td>Somatic symptoms/fatigue (Hamilton depression scale)</td>
<td>-0.02</td>
<td>64 90</td>
<td>14 93</td>
</tr>
<tr>
<td>Decreased libido (Hamilton depression scale)</td>
<td>-0.10</td>
<td>57 80</td>
<td>11 73</td>
</tr>
<tr>
<td>Hypochondriasis (Hamilton depression scale)</td>
<td>-0.12</td>
<td>28 39</td>
<td>4 27</td>
</tr>
<tr>
<td>Insomnia—middle (Hamilton depression scale)</td>
<td>-0.13</td>
<td>39 55</td>
<td>7 47</td>
</tr>
<tr>
<td>Depressed mood (Hamilton depression scale)</td>
<td>-0.24</td>
<td>71 100</td>
<td>15 100</td>
</tr>
<tr>
<td>Social withdrawal (scale for atypical symptoms)</td>
<td>-0.39</td>
<td>67 94</td>
<td>14 93</td>
</tr>
<tr>
<td>Decreased activity (Hamilton depression scale)</td>
<td>-0.41</td>
<td>71 100</td>
<td>15 100</td>
</tr>
<tr>
<td>Combined score on scale for atypical symptoms and Hamilton depression scale</td>
<td>-0.42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guilt (Hamilton depression scale)</td>
<td>-0.48</td>
<td>55 77</td>
<td>14 93</td>
</tr>
<tr>
<td>Loss of appetite (Hamilton depression scale)</td>
<td>-0.48</td>
<td>15 21</td>
<td>6 40</td>
</tr>
<tr>
<td>Psychic anxiety (Hamilton depression scale)</td>
<td>-0.53</td>
<td>60 85</td>
<td>14 93</td>
</tr>
<tr>
<td>Insomnia—early (Hamilton depression scale)</td>
<td>-0.59</td>
<td>15 21</td>
<td>6 40</td>
</tr>
<tr>
<td>Insomnia—late (Hamilton depression scale)</td>
<td>-0.66</td>
<td>19 27</td>
<td>9 60</td>
</tr>
<tr>
<td>Somatic anxiety (Hamilton depression scale)</td>
<td>-0.67</td>
<td>42 59</td>
<td>13 87</td>
</tr>
<tr>
<td>Mornings worse (Hamilton depression scale)</td>
<td>-0.67</td>
<td>30 42</td>
<td>11 73</td>
</tr>
<tr>
<td>Depersonalization (Hamilton depression scale)</td>
<td>-0.76</td>
<td>10 14</td>
<td>7 47</td>
</tr>
<tr>
<td>Suicidality (Hamilton depression scale)</td>
<td>-0.77</td>
<td>23 32</td>
<td>10 67</td>
</tr>
<tr>
<td>Retardation (Hamilton depression scale)</td>
<td>-0.82</td>
<td>12 17</td>
<td>6 40</td>
</tr>
<tr>
<td>Total score on Hamilton scale</td>
<td>-1.18</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*List is arranged in the order of the effect sizes. Positive effect sizes indicate predominance of the symptom among responders, and negative effect sizes indicate predominance of the symptom among nonresponders. Items with at least a moderate effect size (d=0.45) were considered potential predictors of treatment response.

*Score on the scale for atypical symptoms divided by score on combined Hamilton depression scale and scale for atypical symptoms.

Atypical balance scores for responders (mean=47.41%, SD=9.81%, versus mean=36.53%, SD=12.95%) (t=3.68, df=84, p<0.001). All subjects with Hamilton scores below 15 were responders (29 [33.7%] of 86 subjects), regardless of atypical symptom severity. Additionally, all subjects with atypical balance scores below 29% were nonresponders (six [7.0%] of 86 subjects). Thus, Hamilton score and atypical balance in these extreme subgroups accounted for treatment outcome in about 41% of cases.

Effect size of symptoms and scales. To determine whether specific symptoms distinguished responders from nonresponders, mean differences at baseline were calculated as effect sizes (d) for items that occurred with at least 20% frequency. These are presented, along with comparisons of combined scale scores, in Table 1. Positive effect sizes indicate predominance of the symptom among responders, and negative effect sizes indicate predominance of the symptom among nonresponders. Items with at least a moderate effect size (d=0.45) were considered potential predictors of treatment response. All items with a positive effect size were atypical symptoms: hypersomnia, afternoon or evening slump, reverse diurnal mood variation (evenings worse), and carbohydrate craving. Items with large negative effect sizes included retardation, suicidality, depersonalization/derealization, classical diurnal mood variation (mornings worse), somatic anxiety, early and late insomnia, psychic anxiety, appetite loss, and guilt.

Several high-frequency symptoms had relatively small effect sizes and thus did not discriminate between the two groups. In one extreme example, fatigability—seen in more than 90% of the subjects—had an effect size of nearly 0. Similarly, effect sizes were small for depressed mood, social withdrawal, and decreased activity (the Hamilton measure of anhedonia). Of the scale measures, the total score on the Hamilton scale showed the largest effect size (d=1.18), closely followed—with the opposite sign—by the atypical balance score (d=1.03). The total scores on the atypical scale and the combined scale showed smaller, moderate effects.

Discriminant analysis. The statistical associations among the scale items provide additional predictive power. In a forward stepwise discriminant analysis of symptoms that exceeded 20% frequency at baseline, the canonical variable—the linear composite that best discriminated between responders and nonresponders—was Y=−0.24 + 0.36 suicidality + 0.50 retardation + 0.49 late insomnia + 0.49 depersonalization + 0.44 afternoon or evening slump. The equation is dominated by four items on the Hamilton scale, with one atypical symptom with
Figure 1. Tree diagrams illustrating hierarchical clustering of baseline symptoms for subjects with winter depression who did (N=71) or did not (N=15) respond to light treatment.

An opposite sign. Several symptoms that showed large mean differences between responders and nonresponders—e.g., hypersomnia and typical diurnal mood variation—are absent from the equation. By a jackknifed classification, the canonical variable correctly identified 77.4% of responders and 63.4% of nonresponders. An equation based on all symptoms regardless of frequency accounted for 84.9% of responders and 71.4% of nonresponders.

Symptom clusters. Intercorrelations among baseline symptoms may differ between responders and nonresponders. Figure 1 displays the results of hierarchical cluster analyses for items exceeding 20% frequency. Those most closely associated with depressed mood were considered to form a core cluster. The tree diagram for responders reveals two primary clusters: the core cluster and a second cluster. The core cluster includes carbohydrate craving, increased eating, increased appetite, hypersomnia, afternoon or evening slump, depressed mood, decreased activity, social withdrawal, and fatigueability. The three food-related symptoms (carbohydrate craving and increased eating and appetite) form a subcluster within the core. In the second major cluster, weight gain is positioned immediately adjacent to the core. Reverse diurnal variation (evenings worse), which was the third highest ranking positive predictor in the effect size analysis (Table 1), is located away from the core.

By contrast, the diagram for nonresponders shows a core cluster that does not include atypical symptoms other than fatigability (which is ubiquitous in seasonal affective disorder) and social withdrawal. The core cluster for nonresponders also includes decreased libido, guilt, fatigue, psychic and somatic anxiety, depressed mood, and decreased activity—a clinical picture of melancholic depression. Most of the atypical symptoms are grouped in a low-ranking tertiary cluster.

Intercorrelations among the symptoms were also examined for the group as a whole by multidimensional scaling, with similar results (not shown): a majority of the atypical items were grouped, and depressed mood and decreased activity formed an aggregate. Similarly, with a principal component factor analysis, varimax rotation of the first factor contained all but one atypical symptom on one end and a group of melancholic symptoms (guilt, late insomnia, and appetite loss) at the opposite pole.

Degree of Improvement for the Group as a Whole

In addition to discriminating extreme cases (responders versus nonresponders), our broader objective was to predict graded changes in clinical response within the group as whole, including partial responders. Therefore, we investigated the predictive utility of rating scale measures as well as the severity of individual symptoms.

Correlations among the scales. To draw correspondences with previous studies, we compared Spearman correlations between baseline and posttreatment (period 1 for crossovers) rating scale measures for raw scores, linear change scores, percent change scores, and deviations from the regression line. Significant r values, all with df=92, are reported at p<.05.

For raw scores, pre- and posttreatment Hamilton scores showed a significant positive correlation, as did combined scores, indicating that severity at baseline was mirrored in the degree of symptom persistence after treatment. By contrast, pre- and posttreatment atypical
scores did not show a significant correlation. The lower the atypical balance at baseline, the higher was the posttreatment Hamilton score (r=0.31); by contrast, the correlation between atypical balance and posttreatment atypical score was nearly 0.

For linear change, baseline scores on each scale showed significant positive correlation with change scores on the same scale (Hamilton, r=0.29; atypical, r=0.66; combined, r=0.43). Furthermore, baseline scores on the combined scale were significantly correlated with change scores on both the Hamilton (r=0.28) and atypical (r=0.34) scales. Atypical balance was significantly correlated with change on the atypical (r=0.61) and combined (r=0.34) scales but not the Hamilton scale (r=0.06).

For the proportional transform of percent change (21), only the baseline atypical score was correlated with improvement on the same scale (r=0.25). The high correlation found between baseline score on the combined scale and linear change score was lost when percent change was used (r=0.05). Of all the baseline measures, atypical balance provided the most inclusive predictor of percent change, yielding significant positive correlations with all three scales (Hamilton, r=0.20; atypical, r=0.31; combined, r=0.26).

As for deviations from the regression line, according to Lord’s early analysis (22), the “embarrassing relationship” between baseline and change scores is eliminated when change is measured by the residual values after regression of posttreatment on baseline scores. The correlation within scales is thus 0. By this analysis, the only significant predictive relationship was between atypical balance and posttreatment combined score (r=0.23), which represents a modest effect size.

Forward stepwise multiple regression. Unlike the independent correlations just presented, multiple regression accounts for mutual interdependence among the measures. When baseline scores on the Hamilton, atypical, and combined scales, plus atypical balance and CGI scores, were entered as independent variables, with percent improvement in combined scale score as the dependent variable, atypical balance was the sole measure to enter in the following equation: Y=28.55 + 71.94 atypical balance (r=0.27, df=1, 82, p<0.05).

When the individual scale items were entered, and the dependent variable was percent improvement on the combined scale, hypomania was the sole symptom to enter the following equation: Y=51.05 + 7.88 hypomania (r=0.28, df=1, 73, p<0.05). By contrast, an analysis of linear change scores on the combined scale—the procedure of Oren et al. (8)—yielded a far more complex equation in which items on both the Hamilton (decreased libido, weight loss, and paranoia) and atypical (hypomania and carbohydrate craving) scales appeared nondifferentially and with positive signs.

Influence of Pretreatment Expectations on Clinical Response

Given the small number of nonresponders, a question remains about placebo factors that could promote non-specific clinical improvement. Expectation ratings obtained at baseline showed no significant correlation with baseline severity on any of the rating scales. However, expectations were significantly higher among responders than nonresponders (mean=5.44, SD=1.27, versus mean=4.36, SD=1.63) (t=2.32, df=48, p<0.025, two-tailed).

The correlation of expectations with improvement was both scale- and symptom-specific. Expectations were not correlated with percent improvement on the atypical scale (r=0.06, df=57, n.s.). However, there was a significant positive correlation with improvement on the Hamilton scale (r=0.32, df=57, p<0.01, one-tailed), which suggests that the classical symptoms of depression are differentially vulnerable to placebo effects. There were moderate negative correlations between expectations and a distinct subset of Hamilton scale items (r=−0.31 to −0.37): guilt, psychic anxiety, suicidality, somatic symptoms/tiredness, and depressed mood. By contrast, expectations showed minimal correlation with atypical symptoms (r=0.22 ≤ r ≤ 0.18).

A Profile of Nonresponders to Light

The cluster analysis (figure 1) indicated that nonresponders could be characterized by an association of melancholic features. When we combined the baseline scores for appetite loss, late insomnia, diurnal variation (mornings worse), guilt, retardation, and agitation, the mean difference between responders and nonresponders was highly significant (t=−4.08, df=84, p<0.001) and the effect size very large (d=1.20).

We expanded this description by analysis of chart records that encompassed a wide range of historical and clinical factors not assessed by the Hamilton or atypical scales. Each factor was scored positive or negative, and a moderate-to-large effect size of proportions (p>0.45) was used as a criterion for distinguishing nonresponders from responders. Thirteen (87%) of the 15 nonresponders and 42 (59%) of the 71 responders reported a family history of depression (h=0.64); five (33%) of the nonresponders and nine (13%) of the responders noted a suicide or suicide attempt in family members (h=0.50). Seven (47%) of the nonresponders and 12 (17%) of the responders were diagnosed as having bipolar disorder or bipolar disorder not otherwise specified (h=0.66). Eleven (73%) of the nonresponders and 27 (38%) of the responders reported past behavior consistent with hypomania (h=0.73). Seven (47%) of the nonresponders and 17 (24%) of the responders reported a history of panic attacks (h=0.48), although none was diagnosed with panic disorder. Six (40%) of the nonresponders and 11 (15%) of the responders reported a history of thyroid deficiency, for which some received hormone replacement (h=0.56) (during the study period, all were in the normal range). Four (27%) of the nonresponders and two (3%) of the responders had been hospitalized for depression (h=0.75). Twelve (80%) of the nonresponders and 25 (35%) of the responders reported having used psychotropic drugs.
(h=0.94). The fact that a large majority of nonresponders to light had previously discontinued medication suggests that they form a treatment-refractory group.

**DISCUSSION**

Responders and nonresponders to light treatment for winter depression showed different clinical profiles when depressed. Although their depression did not differ in severity, as assessed by the combined scale, nonresponders were distinguished by higher Hamilton scores and lower atypical balance (they had a predominance of melancholic features). They also reported histories of medication, hospitalization, and depression in family members, indicative of greater risk. A predictive relation also was found in graded rating scale measures for the whole group, including partial responders: when the inherent dependency of posttreatment status on initial scores was factored out by computing deviations from the regression on posttreatment combined-scale scores, atypical balance stood as the sole significant predictor of clinical improvement.

Several earlier studies showed that baseline scores on atypical symptom scales (6, 7) or specific atypical symptoms (8-10) correlated with the magnitude of treatment response. One study (6) found a negative correlation between baseline Hamilton score and improvement, which implies that atypical balance would be a positive predictor. Two studies that failed to find predictivity in hypersonia (10, 11)—which ranked as our strongest individual symptom predictor—used subjects with generally fewer atypical symptoms. It should be noted also that atypical scale ratings of hypersonia are fallible; prospective sleep logs do not necessarily demonstrate greater sleep duration (23), and interview assessments may reflect perceived need. One study identified difficulty awakening, rather than hypersonia, as a predictor of response to light (24).

By dividing diurnal mood variation into typical and atypical patterns—mornings or evenings worse—we found opposite results for nonresponders and responders. Past studies have been unable to demonstrate a predictive relation between diurnal variation and response to light (8-11), which may reflect smaller sample sizes with fewer atypical symptoms.

A puzzling item on the atypical scale is social withdrawal. At baseline it occurs with greater than 90% frequency in responders and nonresponders alike, and it stands as a negative predictor of improvement with small-to-moderate effect size. For both groups it falls away from the hierarchical cluster of atypical symptoms and is most closely associated with depressed mood, reduced activity, and fatigability. Fatigability is represented on both Hamilton and atypical scales; it was included in the latter scale to provide additional weight as a prominent symptom of seasonal affective disorder. Indeed, fatigability is ubiquitous among responders (100%) and nonresponders (97%) alike and is greatly reduced with successful light treatment. However, of all symptoms it is the least predictive of treatment success and is most closely associated in a hierarchical cluster with depressed mood and reduced activity rather than with the other atypical symptoms. The placement of social withdrawal and fatigability on the atypical scale might be questioned, and the predictive power of the various scale measures might be enhanced by reassignment to the Hamilton group. Similarly, reverse diurnal mood variation might be appropriately assigned to the atypical scale.

Our data indicate a clear negative predictive relation between suicidality and treatment response. The effect size between responders and nonresponders was large, and suicidality showed the greatest weight of all symptoms in the discriminant analysis, for which the canonical variable was dominated by negative predictors on the Hamilton scale. In a multiple regression analysis, Oren et al. (8) identified suicidality as a positive predictor, which presents an apparent paradox. We infer that their result derived from the use of linear change scores dependent on baseline severity. Thus, even though suicidality declines after treatment, its presence at baseline does not predict treatment success.

Meesters et al. (11) reported superior improvement with greater baseline severity, which contrasts with our results and the conclusions of the cross-center study (2). In a replication test against our preliminary data (25), we found that Meesters et al. also failed to isolate the atypical balance factor. However, they used raw scale scores with inherent dependency on the baseline, an atypical scale lacking the item for afternoon or evening slump (which was the sole atypical predictor in our discriminant analysis), and a small number of subjects with generally few atypical symptoms and low relapse rates.

We found a trend toward a higher proportion of responders among unipolar than bipolar patients (72.9% [N=62] versus 50.0% [N=9]) (χ²=3.65, df=1, p=0.06), similar to the results of Stinson and Thompson (6). We found only a small negative correlation between age and percent improvement (rs=-0.11, df=92, n.s.), which does not replicate Lam’s significant result using multiple regression on linear change scores (9). However, as in Lam’s study, we found no sex difference in treatment response.

Two studies of light treatment for seasonal affective disorder included several failures to distinguish clinical response from placebo response (26, 27). Our finding of a selective negative correlation between pretreatment expectations and improvement in Hamilton score suggests the presence of a nonspecific effect on mood-related symptoms. The contrast between the Hamilton and atypical scales argues against a simple expectation-based placebo explanation of overall response and is consistent with the hypothesis that light exerts a specific, active effect on atypical vegetative symptoms in seasonal affective disorder. Lower expectations among nonresponders might reflect pessimism stemming from difficulties with past treatment (which they received more frequently than did responders) or greater severity of typical depressive symptoms.
Given the lack of treatment response and the melancholic profile as guideposts, the question arises as to whether nonresponders should be diagnosed as having seasonal affective disorder. Detection of seasonality is fallible when current criteria are used. DSM-III-R requires at least three winter episodes, two of which were consecutive, and a ratio of seasonal-to-nonseasonal episodes greater than three-to-one (a criterion that has been diluted in DSM-IV to a "regular relationship" between episodes and season). NIMH criteria for seasonal affective disorder require two consecutive winter episodes. Both sets of criteria allow for a variable history. Indeed, there have been prospective reports of seasonal depression that became nonseasonal depression and vice versa (28). Despite a high frequency of atypical symptoms, subjects with nonseasonal atypical depression show minimal improvement in response to light (29). Two longitudinal studies showed that only 22% (28) and 35% (30) of patients with initial diagnosis of seasonal affective disorder subsequently exhibited reliable winter recurrence. However, a retrospective study based on monthly interview data showed an annual risk factor of about 70% in patients with seasonal mood disorders (31). Consideration of the symptom profile may add specificity to the diagnosis of seasonal affective disorder, especially given the potential variability in seasonal pattern. Indeed, atypical vegetative symptoms are more prevalent in patients who show stable seasonality (28). This longitudinal observation converges on our finding that atypical balance predicts the response to light, which in turn might be considered a defining feature of the syndrome.

REFERENCES

24. Avery D, Bolte MA, Eder D: Difficulty awakening as a symptom of winter depression, in Society for Light Treatment and Biological Rhythms Abstracts 1994; 621