Bright Light Treatment of Winter Depression

A Placebo-Controlled Trial

Charmane I. Eastman, PhD; Michael A. Young, PhD; Louis F. Fogg, PhD; Liwen Liu, PhD; Patricia M. Meaden, PhD

Background: Bright light therapy is the recommended treatment for winter seasonal affective disorder (SAD). However, the studies with the best placebo controls have not been able to demonstrate that light treatment has a benefit beyond its placebo effect.

Methods: Ninety-six patients with SAD completed the study. Patients were randomly assigned to 1 of 3 treatments for 4 weeks, each 1.5 hours per day: morning light (average start time about 6 AM), evening light (average start about 9 PM), or morning placebo (average start about 6 AM). The bright light (6000 lux) was produced by light boxes, and the placebos were sham negative-ion generators. Depression ratings using the Structured Interview Guide for the Hamilton Depression Rating Scale, SAD version (SIGH-SAD) were performed weekly.

Results: There were no differences among the 3 groups in expectation ratings or mean depression scores after 4 weeks of treatment. However, strict response criteria revealed statistically significant differences; after 3 weeks of treatment morning light produced more of the complete or almost complete remissions than placebo. By 1 criterion (24-item SIGH-SAD score ≤50% of baseline and ≤8), 61% of the patients responded to morning light, 50% to evening light, and 32% to placebo after 4 weeks of treatment.

Conclusions: Bright light therapy had a specific antidepressant effect beyond its placebo effect, but it took at least 3 weeks for a significant effect to develop. The benefit of light over placebo was in producing more of the full remissions.

Arch Gen Psychiatry. 1998;55:883-889

The shorter photoperiod and decrease in sunlight exposure experienced by people living in temperate and higher latitudes during the winter is hypothesized to be the main trigger for winter depression or seasonal affective disorder (SAD). Supplementation with bright artificial light is therefore logical and is also the medically recommended treatment of choice. There is no doubt that bright light therapy can reduce and even eliminate the symptoms of winter SAD. However, it has remained difficult to demonstrate that bright light has an antidepressant effect beyond its placebo effect. One reason stems from the problems inherent in finding an appropriate placebo control for light box treatment.

Many researchers used dim light as the placebo control for bright light. However, even in the earliest study, most patients predicted that bright light would be more helpful. Because expectations for improvement are believed to account for a large component of placebo effects and placebo effects account for a large component of antidepressant effects, the superiority of bright light could have been entirely due to placebo effects. Another strategy has been to use evening light as the placebo control for morning light. However, some investigators have shown that evening light can be as antidepressant as morning light, limiting the usefulness of evening light as an inert placebo.

See also pages 861, 863, 875, and 890

A recent study successfully implemented an inert placebo control, a light box that emitted no visible light, but had a pilot light and a hum when turned “on.” Patients were told that the purpose of the experiment was to study infrared, nonvisible light. The deception was successful; there was no difference in expectations between the patients assigned to bright light and those assigned to placebo. Unfortunately, there was also no difference in antidepressant response.

Several years ago, we used an inert (deactivated) negative ion generator as the placebo control for bright light treat-
PATIENTS AND METHODS

PATIENTS

Patients were recruited through advertisements and the local media. They were diagnosed by the usual criteria for SAD, but in addition we required the “atypical” symptoms of increased appetite or weight and increased sleep. Most SAD patients have these symptoms, and this provided a more homogeneous and more typical sample. Patients also had to score 21 or more on the 24 items of the Structured Interview Guide for the Hamilton Depression Rating Scale, SAD Version (SIGH-SAD), which is the original 17-item Hamilton Depression Scale plus 7 atypical items, the 24 items that were intended to rate severity. All patients were free of psychotropic medications for several months and none had previously tried bright light or negative ion treatment. None had any complicating medical conditions based on medical histories, physical examinations, and blood and urine evaluations. Patient characteristics are listed in Table 1.

EQUIPMENT

The ion generators were shiny black cylinders, 32 cm in height, 15 cm in diameter, with 3 small lights on the front that changed rapidly between red and green. Because the negative ion mechanism made a soft hissing sound, a white-noise generator was added to active and placebo generators. The patient consent forms explained that they might receive active or deactivated generators, but that they would not be able to tell which type they had received. Two generators were set up on a desk or table in front of the patient, 38.1 cm apart, with 38.1 cm between the patient and each generator.

The light box was 65 cm wide and 43.5 cm tall (Apollo Light Systems, Orem, Utah). It contained 6 horizontally mounted cool-white fluorescent lamps. Patients sat at a table or desk with the light box directly in front of them at a distance of 38.1 cm. They usually read during the treatment time, which produced an illuminance of about 6000 lux.

PROCEDURES

The 5-week protocol consisted of a baseline week followed by 4 consecutive treatment weeks. Patients were told that they would be randomly assigned to 1 of 4 treatments: morning generator, evening generator, morning light, or evening light. All treatments were 1.5 hours in duration. The morning treatments were taken as soon as possible after waking and the evening treatments were before bed, with 1 hour maximum between the end of treatment and bedtime.

Sleep schedules were negotiated with each patient, and were followed during the 5-week study, including weekends. The wake-up time had to be earlier than usual, to make time for the morning treatments. Even patients assigned to evening treatments had to wake up earlier because they were asked to stay indoors for the first 1.5 hours after waking. This was to help us find a difference between morning and evening treatments, if it existed. Bedtime fell within a 1-hour window so that patients were allowed 7 to 8 hours in bed. Naps were permitted within a 6-hour window in the middle of the waking period. Compliance with the sleep schedule was encouraged and monitored by requiring that patients call a time-stamp answering machine or voice mail 3 times each morning, during the 1.5 hours after waking, and by reviewing daily sleep logs at least once per week. Patients were not permitted to drink alcohol during the 5 weeks or to drink caffeine in the 6 hours before their earliest bedtime.

Patients went to 2 different buildings during weekly visits to the medical center. The clinical staff in the Department of Psychiatry, Rush-Presbyterian-St Luke’s Medical Center, Chicago, Ill, performed weekly SIGH-SAD ratings, and were blind to the patients’ type of equipment and time of day of treatment. The research assistants in the Biological Rhythms Research Laboratory dispensed equipment, collected questionnaires, and monitored compliance to the protocol. They were not blind to type of equipment or time of day of treatment, but they did not know which patients received active generators and which received placebos. This division of functions between the 2 locations helped to maintain the blinding of the SIGH-SAD raters.

Treatment was administered 6 days per week; it was omitted on the weekly visit day. Patients started the protocol in January or February, except for 3 who started in November.

Before the study began patients were given a packet of information from the scientific and lay literature promoting negative ion and light treatment of SAD. At the end of the baseline week, after patients had known their treatment assignment for 1 week, they completed an expectation questionnaire that asked them to rate, on a 7-point scale, how they expected to feel at the end of the 4-week treatment. Possible answers ranged from “will feel as good as summer” (1) to “no change” (7). They also rated how confident they were in that guess, from “very confident” (1) to “not at all confident” (7).

At the end of each week patients completed the 21-item Beck Depression Inventory, to which we added 4 atypical items. The 4 items covered increases in appetite, weight, sleep, and sleepiness.

The study protocol was approved by the institutional review board and all patients gave written informed consent before the study began.

DATA ANALYSIS

We analyzed data from patients assigned to morning light (ML), morning placebo (MP), and evening light (EL). Patients and raters did not know that very few patients were assigned to active ion generators and none were assigned to evening generators. Patients were randomly assigned to the 3 major groups with balancing for sex. Statistical analyses included repeated-measures analyses of variance (ANO瓦s) and log-linear analyses for depression scores, and 1-way ANOVAs and Tukey honestly significant difference post-hoc tests for expectation ratings and sleep parameters. In all statistical tests we used a significance level of .05, 2 tailed, unless otherwise specified. Summary statistics are presented as mean ± SD.
Despite this slight advantage for light, the benefit of light treatment over placebo was not statistically significant.

For the current study, we improved the design to increase the chances of finding a significant effect of bright light treatment, should it exist. We increased the daily dose of light treatment from 2 to 4. We coached the staff to be enthusiastic and positive about each patient 2 units instead of 1.

Expectations vs SIGH-SAD difference from pre-baseline to treatment; at treatment week 3, \( r = -0.18, P = .04; \) at treatment week 4, \( r = -0.14, P = .09; \) both with \( df = 94, 1 \)-tailed tests).

### RESULTS

During the course of 6 winters, 96 patients completed the study (Table 1) and 25 dropped out during the 5-week protocol (ML = 8, EL = 8, MP = 9).

### EXPECTATIONS

The expectation ratings showed that all groups expected their treatment to be very successful in reducing their depression (ML = 3.6 ± 1.1, MP = 3.7 ± 1.2, EL = 3.3 ± 1.1). All groups were somewhat confident in their prediction (ML = 3.7 ± 1.4, MP = 3.9 ± 1.3, EL = 3.2 ± 1.2). There were no statistically significant differences among the groups for either item, determined by 1-way ANOVAs. There were small correlations between expectations and depression ratings, showing a trend for patients who expected a greater improvement to achieve a greater reduction in their depression (ex-

### SLEEP

Sleep logs were analyzed to determine whether there were differences among groups (Table 2). If there were differences in sleep schedules or other sleep parameters between light and placebo treatments, then these nonspecific factors might contribute to any differences in antidepressant response. One-way ANOVAs were used to compare treatment groups. Within-group comparisons are only descriptive. There were no statistically significant differences among groups in the scheduled wake-up times, which averaged about 6 AM. Actual wake-up time was about an hour or 2 earlier during the study (baseline and treatment) than before the study (pre-baseline), when there were no restrictions on sleep. However, sleep onset times did not change as much as wake-up times. Consequently, nighttime sleep was about an hour less during the study. Even when naps were added to nighttime sleep to yield total sleep time, sleep during the study was still about an hour less than before. The only significant difference between groups was that the ML group reported slightly more sleep during the 4 treatment weeks compared with the EL group (\( = 20 \) minutes per day more). There were no significant differences in sleep between the ML and MP groups.

## Table 1. Demographic and Clinical Characteristics of Patients*

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Morning Light</th>
<th>Morning Placebo</th>
<th>Evening Light</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of patients (M/F)</td>
<td>33 (6/28)</td>
<td>31 (3/28)</td>
<td>32 (5/27)</td>
</tr>
<tr>
<td>Age, y</td>
<td>35.5 (10.7)</td>
<td>37.0 (9.2)</td>
<td>37.7 (11.3)</td>
</tr>
<tr>
<td>Race/ethnicity, %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>28</td>
<td>28</td>
<td>31</td>
</tr>
<tr>
<td>Black</td>
<td>4</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Asian</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Hispanic</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>% Married</td>
<td>57.6</td>
<td>45.2</td>
<td>34.4</td>
</tr>
<tr>
<td>Education, y</td>
<td>16.2 (2.4)</td>
<td>16.1 (2.5)</td>
<td>16.7 (2.4)</td>
</tr>
<tr>
<td>Duration of SAD, y</td>
<td>14.3 (7.9)</td>
<td>13.1 (6.8)</td>
<td>14.8 (8.0)</td>
</tr>
<tr>
<td>Age at onset of SAD, y</td>
<td>22.1 (9.8)</td>
<td>22.8 (7.7)</td>
<td>21.7 (10.1)</td>
</tr>
<tr>
<td>Previous treatment, %</td>
<td>36.4</td>
<td>29.0</td>
<td>34.4</td>
</tr>
<tr>
<td>Medication</td>
<td>51.5</td>
<td>48.4</td>
<td>65.6</td>
</tr>
<tr>
<td>Psychotherapy</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Data are given as mean (SD) unless otherwise indicated. SAD indicates seasonal affective disorder.

## Table 2. Sleep Parameters*

<table>
<thead>
<tr>
<th></th>
<th>Morning Light</th>
<th>Morning Placebo</th>
<th>Evening Light</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scheduled time†</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wake-up</td>
<td>6.1 (1.3)</td>
<td>6.2 (1.1)</td>
<td>6.4 (0.9)</td>
</tr>
<tr>
<td>Latest bedtime‡</td>
<td>23.1</td>
<td>23.2</td>
<td>23.4</td>
</tr>
<tr>
<td>Sleep onset†</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-baseline week</td>
<td>7.9 (1.4)</td>
<td>7.9 (1.2)</td>
<td>7.6 (1.1)</td>
</tr>
<tr>
<td>Baseline week</td>
<td>6.2 (1.3)</td>
<td>6.1 (1.2)</td>
<td>6.5 (1.0)</td>
</tr>
<tr>
<td>Treatment weeks§</td>
<td>6.1 (1.3)</td>
<td>6.1 (1.2)</td>
<td>6.5 (0.9)</td>
</tr>
<tr>
<td>Time (in hours) from sleep onset to wake-up minus any awakenings of 10 minutes or more.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nighttime sleep¶</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-baseline week</td>
<td>7.8 (0.9)</td>
<td>8.0 (0.8)</td>
<td>7.9 (1.2)</td>
</tr>
<tr>
<td>Baseline week</td>
<td>6.7 (0.6)</td>
<td>6.6 (0.5)</td>
<td>6.8 (0.5)</td>
</tr>
<tr>
<td>Treatment weeks§</td>
<td>6.9# (0.4)</td>
<td>6.7 (0.5)</td>
<td>6.6 (0.4)</td>
</tr>
<tr>
<td>Total sleep time¶</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-baseline week</td>
<td>8.3 (0.9)</td>
<td>8.3 (1.0)</td>
<td>8.1 (1.2)</td>
</tr>
<tr>
<td>Baseline week</td>
<td>7.0 (0.7)</td>
<td>6.9 (0.6)</td>
<td>7.0 (0.7)</td>
</tr>
<tr>
<td>Treatment weeks§</td>
<td>7.2# (0.5)</td>
<td>7.0 (0.6)</td>
<td>6.9 (0.5)</td>
</tr>
</tbody>
</table>

*All data are given as mean (SD).
†Hours past midnight.
‡Seven hours before scheduled wake-up time.
§Mean of all 4 treatment weeks.
||Time (in hours) from sleep onset to wake-up minus any awakenings of 10 minutes or more.
|Nighttime sleep plus naps.
#Significantly different from evening light (P < .05).
also not significant (F6,279 = 1.01, \( P < .05 \)) and not of treatment group (F2,93 = 0.28, \( P = .75 \)). The interaction of group \( \times \) time was also not significant (F6,279 = 1.01, \( P = .42 \)). Thus, patients improved over time regardless of treatment group. Two tests, the Mauchly sphericity test and the Box M, both revealed that the assumption of homogeneous variances was violated. Therefore, it would not be advisable to rely solely on the ANOVAs for statistical significance.

We classified patients as responders if their 24-item SIGH-SAD score decreased to 50% of baseline. The percent responders for the ML, EL, and MP groups were 55%, 50%, and 32%, respectively. Morning light was significantly better than placebo, but the difference did not reach statistical significance until the third week of treatment. A lag of 3 weeks emerged when we classified patients as responders by strict joint criteria designed to identify those with complete or nearly complete remissions (a decrease in SIGH-SAD of 50% and a score \( \leq 8 \)), criteria similar to those used previously.21 Figure 2 shows that the percent responders increased as the weeks progressed, and that ML produced the best response rates. At treatment week 4, the percent responders for the ML, EL, and MP groups were 61%, 50%, and 32%, respectively. Log-linear analyses revealed significant differences among treatments; contrast parameters showed that the response rate for ML was higher than for MP at week 3 (\( z = 2.55, P < .05 \)) and at week 4 (\( z = 2.24, P < .05 \)). In addition, there were more responders to ML than EL at each treatment week, but this difference was only significant during treatment week 3 (\( z = 2.13, P < .05 \)). We also analyzed these data by applying even stricter response criteria. Responders had to meet the same criteria for complete remission as mentioned in Figure 2, but they had to meet these criteria in both treatment weeks 3 and 4. By these criteria the response rates for the ML, EL, and MP groups were 55%, 28%, and 16%, respectively. Morning light was significantly better than MP (\( z = 3.05, P < .01 \)) and significantly better than EL (\( z = 2.13, P < .05 \)).

The reason that significant differences between groups emerged when calculating percent responders (Figure 2), but not with mean ratings (Figure 1) can be understood by considering the distributions of scores (Figure 3). The scatterplots show that more patients met the strict joint criteria for remission in the ML group than in the MP or EL groups. The frequency histograms further illustrate how the scores were distributed differently in the 3 groups. The ML distribution was Poisson-like, with scores clustered at the bottom, whereas the MP distribution was more normal. In summary, although the mean scores were similar among the 3 groups (\( = 10-12 \)), ML produced the greatest number of full remissions.

The results from the Beck depression ratings were similar to the SIGH-SAD ratings. There was little difference among the 3 groups in mean weekly ratings. Log-linear analyses on percent responders by strict criteria showed significantly higher response rates to ML than MP. However, there were no significant differences between ML and EL (Table 3).

Bright light therapy was a better antidepressant for SAD than placebo, but the difference did not reach statistical significance until the third week of treatment. A lag of
at least 3 weeks for significant antidepressant effects to develop is consistent with antidepressant drug studies. Thus, it seems likely that the reason our previous efficacy study and the other light box study with a good placebo control (no light, infrared deception) could not demonstrate a difference between light and placebo was because the treatment time was too short. In both cases, the light treatment was only 2 weeks. The infrared deception study also included head-mounted light visors, but there was no difference in antidepressant response between light and no-light visors. There have been 4 other visor studies, all with 1- or 2-week treatments, that did not show a superiority for the presumed active visor despite large sample sizes. These studies used parallel designs, so that patients only saw their own visor and could not tell if it was a placebo. We believe that these 5 studies are the best efficacy studies of light therapy published so far, because data showed that patients did not have higher expectations for the presumed active treatment than for the placebo.

However, there have been several SAD light box studies that found bright light more antidepressant than dim light, sometimes dramatically better after only 1 or 2 weeks. It is possible that those differences were primarily produced by differences in expectation/placebo effects because these were light box studies, mostly crossover designs, in which the bright vs dim comparison was obvious. This explanation is supported by an analysis of light box studies showing that responders had higher expectations than nonresponders. Furthermore, it is well known that the apparent efficacy of antidepressant treatments decreases as designs are made more sophisticated; eg, with more adequate control treatments and better blinding techniques.

It has been suggested that a certain amount of positive expectations or placebo effect may be necessary to release the specific pharmacologic effect of a drug or the specific effects of other treatments. This is one reason we strived to increase the positive expectations of our patients. Thus, as expected, there was a larger placebo response rate in the current study than in our previous study: 36% vs 22% of patients were responders with placebo (<50% of baseline with 2 weeks of MP). Similarly, there was a larger response rate to bright light, 49% vs 34% (with 2 weeks of ML). The increase in response rate for light was similar to the increase for placebo. Thus, the increased response rate to light in the current study can be accounted for by the increased placebo component rather than factors specific to light treatment, such as the increase in daily dose from 1.0 to 1.5 hours. Conversely, the low response rate to light in the previous study can be attributed to a smaller placebo component, rather than a deficient dose of light.

In this study, the largest component of the antidepressant response to bright light was the placebo effect, as in antidepressant drug studies. The placebo effect is composed of many factors, including anxiety reduction associated with receiving a diagnosis and positive expectations for treatment. In our study, another factor was the regular sleep schedule, which could have functioned as a zeitgeber to stabilize circadian rhythms and improve sleep. The restrictions on drinking alcohol and caffeinated beverages also constitute basic sleep hygiene and may have improved sleep. On the other hand, the strict early sleep schedule produced chronic partial sleep deprivation, which could have exacerbated symptoms, especially fatigue, or conceivably produced an antidepressant effect. Thus, there are many non-specific factors that could have contributed to the placebo effect. The fact that we found only small correlations between expectation ratings and antidepressant response makes sense given that patient expectations...
were only one of the factors influencing antidepressant response. The important principle for an efficacy study is to make sure that all the nonspecific factors, including expectations, are the same in both the active and placebo conditions.

Because we succeeded in creating a placebo treatment that was similar to light treatment in many ways (an electrical device that creates an environmental factor, 1.5 hours of regular sitting time, the early sleep schedule, and so on), and because expectation ratings showed no difference between light and placebo, we are fairly confident that bright light had a specific antidepressant effect beyond its placebo effect. However, our placebo could be criticized for being obviously different from a light box. Our expectation questionnaire, although purposely given by research assistants (not by more formidable PhDs) and without any pressure for specific responses, still may not have captured the true feelings of our patients. If they actually had higher expectations for light, that could account for the superior antidepressant effect of light in our study. On the other hand, our design had an advantage over the typical double-blind antidepressant drug studies, in which raters can often tell which patients were given light by the emergence of side effects.13,33 Our raters had no similar clues to guess which patients had light boxes and which had negative-ion generators.

There has been a controversy about whether ML is a better antidepressant than EL, with some studies showing a superiority for ML,14,41-44 but others showing no difference between ML and EL.15,45-48 We found that ML was better than EL by some measures but not by others; the difference between ML and EL was not very robust. Those differences we did find cannot be attributed to differences in expectations, because these 2 groups had similar expectations. However, patients receiving EL obtained slightly less sleep during the treatment weeks than patients receiving ML. Thus, ML could have produced a slightly better antidepressant response because it advances circadian rhythms and can help patients adapt to an earlier sleep schedule, whereas EL delays rhythms and can make it more difficult to adjust.49 The ML and EL groups also differed in the restrictions on social activities. Patients receiving EL had to stay at home (average time 9:00-10:30 PM) to receive their treatment. This interfered with many social and family activities and could have worked against an antidepressant effect. If we had included an evening placebo group, we might have been able to learn more.

Accepted for publication March 24, 1998.

This research was supported by grant MH42768 from the National Institutes of Mental Health, Rockville, Md.


Eva Chernin and Patricia Meaden, PhD, were research coordinators in the Department of Psychiatry, Rush-Presbyterian-St Luke's Medical Center, Chicago, Ill. Other staff in the Department of Psychiatry who helped with recruitment, diagnoses, and SIGH-SAD ratings were James Murakami, Marilyn Perlman, Kay Peterson, Jill M. Richards, and Kelly J. Sherack. William Schefter, MD, reviewed the laboratory test and physical examination reports made during the screening process. Research assistants in the Biological Rhythms Research Laboratory of the Department of Psychology who saw patients and entered data into the computer were David Anyadike, Linda C. Gallo, Christina Hardway, Erin K. Hoese, Michael P. Mahoney, Paula J. Mitchell, PhD, Trina Moss, Charles Splete, and Robert A. Tell. Larry D. Chait, PhD, suggested using negative ion generators as the placebo treatment. Ron Rehr of the Machine Shop and Phillip Zervos of the Electronic Research Facility of Rush-Presbyterian-St Luke's Medical Center helped design and build the negative ion generators.

Reprints: Charmaine I. Eastman, PhD, Biological Rhythms Research Laboratory, Rush-Presbyterian-St Luke's Medical Center, 1653 W Congress Plkwy, Chicago, IL 60612 (e-mail: ceastman@rush.edu).

REFERENCES

1. Eastman CI. Natural summer and winter sunlight exposure patterns in seasonal affective disorder. Physiol Behav. 1990;48:611-616.
12. Eastman CI. What the placebo literature can tell us about light therapy for SAD. Psychopharmacol Bull. 1990;26:495-504.


