# Light Therapy for Seasonal Affective Disorder

## A Review of Efficacy

*Michael Terman, Ph.D., Jiuan S. Terman, Ph.D., Frederic M. Quitkin, M.D., Patrick J. McGrath, M.D., Jonathan W. Stewart, M.D., and Brian Rafferty, A.B.*

*Bright artificial light has been found effective in reducing winter depressive symptoms of Seasonal Affective Disorder, although conclusions about the true magnitude of treatment effect and importance of time of day of light exposure have been limited by methodologic problems. Individual subjects' data from 14 research centers studying 332 patients over 5 years were analyzed with a pooled clustering technique. Overall, 2500-lux intensity light exposure for at least 2 hours daily for 1 week resulted in significantly more remissions—Hamilton Depression Rating Scale (HAM-D) score reduction of 50% or more to a level under 8—when administered in the early morning (53%) than in the evening (38%) or at midday (32%). All three times were significantly more effective than dim light controls (11%). Dual daily exposures (morning-plus-evening light) provided no benefit over morning light alone. In morning–evening crossovers, remission rates were 62% under morning light alone, compared with 28% under evening light alone, with a differential morning–evening response present in 59% of morning responders compared with 10% of evening responders ($p < 0.001$). Remission rates with morning light were highest given low severity at baseline (HAM-D score of 10–16: 67% remission), as compared with moderate-to-severe cases (HAM-D score above 16: approximately 40% remission) where no morning–evening differences were found. Firmer conclusions await treatment studies with larger sample sizes and full assessment of atypical vegetative symptoms seen in winter depression but underrepresented in the Hamilton scale. Longer treatment course and greater light intensity may help clarify clinical response despite the impossibility of achieving a conventional blind placebo control.* ***[Neuropsychopharmacology 2:1–22, 1989]***

KEY WORDS: *Affective disorders; Phototherapy; Circadian rhythms; Antidepressive agents; Placebos; Clinical trials*

A syndrome of annually recurring clinical depressions, which occur in the late fall and winter within the temperate zones, has recently been identified and has prompted novel therapeutic experiments involving administration of indoor supplements of artificial light. Reports during the last 5 years that this simple nonpharmacologic intervention can induce complete clinical remissions within a few days, matching or surpassing the effects of antidepressant medications, prompt a critical review at this point. Seasonal affective disorder (SAD; Rosenthal et al. 1984) is distinguished from other depressive syndromes by clinical course and symptom profile. Common symptoms include sadness, anxiety, irritability, premenstrual difficulties, and decreased energy, activity, and libido. In addition, SAD is marked by prominent atypical vegetative changes including hypersomnia, increased appetite (especially for carbohydrates) and weight, and difficulties with work and interpersonal function.

Reports of symptoms increase with distance from the equator and in climatic conditions of low light

From the Department of Psychiatry, Columbia University (M.T.; F.M.Q.; P.J.M.; J.W.S.) and New York State Psychiatric Institute (M.T.; J.S.T.; F.M.Q.; P.J.M.; J.W.S.; B.R.).

Address reprint requests to: Dr. Michael Terman, Box 50, NYS Psychiatric Institute, 722 West 168th Street, New York, NY 10032.

Received August 1, 1988; revised November 18, 1988; accepted November 20, 1988.


Published by Elsevier Science Publishing Co., Inc.
655 Avenue of the Americas, New York, NY 10010

0893-133X/89/$3.50

(Lingjaerde et al. 1986; Potkin et al. 1986). A questionnaire survey of a New York City random sample indicated that up to 25% of the population suffers from this condition, whether at clinical or subsyndromal levels (Terman 1988a). An extension of the survey to communities spanning the eastern United States indicated strong latitude dependency, a tenfold increase in clinical severity from Sarasota, Florida, to Nashua, New Hampshire (Rosenthal et al. 1988c). SAD patients often report rapid remission of symptoms when visiting locations between the equator and 30°N latitude.

Demographic and clinical features of SAD have been summarized for large groups of diagnosed patients in Bethesda, Maryland ($n$ = 246; cf. Hellekson 1989), London ($n$ = 51; Thompson and Isaacs 1988), New York City ($n$ = 163; Terman et al. 1989a), and Switzerland ($n$ = 63; cf. Hellekson 1989). Mean age of onset usually occurs in the early 20s. Female patients outnumber males by 4:1 or greater. Onset of depressive symptoms typically occurs after the equinox in autumn; the most difficult months are January and February; and spontaneous remissions typically occur after the vernal equinox. The Research Diagnostic Criteria (Spitzer et al. 1978) shows SAD patients with Unipolar, Bipolar I and II, and Minor Depression, with considerable variation in the proportions across centers. Family history of affective disorder is common (25% to 58% across centers), with frequent reports of alcohol abuse or dependence (8% to 36%) and a SAD symptom pattern (14% to 23%). Many patients never used antidepressant medications (26% to 58%), though 66% of the New York group, for example, sought psychotherapy.

Since the initial demonstration of antidepressant effect of bright artificial light with a bipolar SAD patient (Lewy et al. 1982), many investigators have replicated the finding, based on a procedural framework developed at the National Institute of Mental Health (Rosenthal et al. 1984, 1985a). Treatment has generally involved daily administration of bright artificial light indoors, most often at home within outpatient protocols. Most studies have used full-spectrum fluorescent light of approximately 2500 lux illuminance at eye-level (at a distance of about 1 m), an intensity obtained outdoors within a few minutes of sunrise. Individual reports, however, have also asserted positive results for incandescent light (Yerevanian et al. 1986) and cool-white fluorescent light (Lewy et al. 1987). Use of dim fluorescent light of 400 lux or lower (red, yellow, and full-spectrum) generally has been construed as an inactive control, although several investigators have suggested that some antidepressant effect is present (Wirz-Justice et al. 1986; Isaacs et al. 1988). In most studies, patients were exposed to light for 2 to 6 hours per day, although some patients have apparently benefited from durations of 30 minutes to 1 hour. The antidepressant response usually occurs within 3 to 4 days of treatment, with a similar time course for relapse during withdrawal. A few studies, however, have suggested protracted posttreatment benefit of up to several weeks (Yerevanian et al. 1986; C. Eastman, personal communication; A. Wirz-Justice, personal communication).

The specific actions underlying the therapeutic effects of bright light exposure have not yet been clarified, and the pathophysiology of SAD is under active investigation. The NIMH group has performed extensive studies on possible biologic substrates (for a review, see Skwerer et al. 1988). Plasma norepinephrine, measured in a challenge test, appears potentiated under light therapy. Immune response of peripheral blood lymphocytes, stimulated by mitogens, is normalized from a high baseline in SAD patients. Long-latency visual-evoked response, P300, which reflects attention factors, was magnified under light therapy in NIMH tests, although work by Murphy et al. (1989) did not obtain a significant change. SAD patients showed higher REM density (but not shorter latency) than normals, with increases in delta sleep and sleep efficiency after light treatment. A wide range of biochemical factors (growth hormone, thyroid-stimulating hormone, and the thyroid hormones, prolactin, cortisol, and melatonin) have been monitored in patients when depressed and in remission, and in comparison to controls, but SAD-specific functions have been elusive. Wurtman and Wurtman (1989) have proposed that serotonin dysregulation in SAD patients may result in reduced wintertime availability of this neurotransmitter: direct serotoninergic stimulation with *d*-fenfluramine has shown a therapeutic benefit. It remains to be specified how light would act to increase serotonin availability and whether this would involve circadian specificity. Bright-light supplements may satisfy a general, magnified light "need" during the dark months of the year, as in the photon-counting hypothesis of Rosenthal et al. (1985b). If so, light therapy should be effective at any time of day, though possibly its effectiveness would be modulated by diurnal variation in the eye's sensitivity (cf. Terman and Terman 1985; Bassi and Powers 1986).

A major unresolved issue is whether the time of day of light exposure is critical. Several studies suggest that it is not, on the basis of nondifferential responses to morning, midday, or evening schedules (James et al. 1985; Hellekson and Rosenthal 1986; Hellekson et al. 1986; Wehr et al. 1986a,b; Jacobsen et al. 1987; Terman et al. 1987; Wirz-Justice et al. 1987). In other studies, morning light has been effective but evening light ineffective or less effective (Lewy et al. 1987; Avery et al. 1988; Terman et al. 1989a). The

question of timing has important theoretical implications for understanding the specific action of light underlying the therapeutic effect. A circadian rhythm mechanism would be implicated if time of day were critical (Lewy and Sack 1986). Both animal (e.g., DeCoursey 1960) and human (e.g., Honma et al. 1987) temporal isolation studies show phase shifts of circadian rhythms that depend on the timing of light presentation.

It is possible that the circadian rhythms of SAD patients are abnormally delayed in winter (Lewy et al. 1987) and that morning light exposure selectively provides a corrective phase advance. Indeed, phase advances of the nocturnal melatonin secretion rhythm (Lewy et al. 1987; Terman et al. 1988a) and of body temperature rhythm (Avery et al. 1988) have been shown to accompany therapeutic responses to morning light. Normals also show wintertime delays in circadian rhythms (Bojkowski and Arendt 1988), although possibly less extreme than those of patients and with smaller phase responses to light administered at a standard morning hour (Lewy et al. 1987). Subsyndromal SAD is common in the general population (Terman 1988a), however, and is responsive to bright light treatment (Kasper et al. 1989), making it difficult to construe "normals" as a straightforward control group for patients.

On the basis of several overnight studies, Wehr et al. (1986a) concluded that melatonin phase advances need not occur under effective phototherapy. Rosenthal et al. (1987) reported no consistent differences in melatonin patterns under bright light phototherapy and a dim light control. In both cases, however, the lack of effect could be attributed to the measurement protocol, in which potentially suppressive bright evening light was presented in close association with expected melatonin onset. Dim light at normal room levels is generally less active than light at outdoor daylight levels in suppressing nocturnal melatonin secretion (Lewy et al. 1980), but there is wide interindividual variation in sensitivity (Bojkowski et al. 1987; Strassman et al. 1987; Brainard et al. 1989) and a graded dose dependency. Nonseasonal bipolars have shown supersensitivity to light in this respect (Lewy et al. 1981). SAD patients may not show such supersensitivity despite their clinical response to supplementary bright light in winter (Murphy et al. 1989; A.J. Lewy, personal communication). Regardless of whether circadian phase responses prove to underlie the treatment effect, suppression of melatonin secretion in itself does not appear to be directly antidepressant (Rosenthal et al. 1985b; Wehr et al. 1986a; Rosenthal et al. 1988a).

Before the issues of timing and mechanism of action can be resolved, however, the clinical data collected from individual light therapy studies require critical evaluation. Problems include small samples, unconvincing controls, lack of blind assessment, lack of corroborative global ratings, variable sleep schedules, and uncontrolled light exposure of outpatients in their living and working environments. Although intake screening has been modeled on criteria for SAD originally developed by Rosenthal et al. (1985a), patients have varied with respect to the presence of atypical vegetative symptoms, concurrent use of medications, and overall severity of symptoms. Until procedural discrepancies are resolved and the significance of these other issues determined, conclusions derived to date must be considered tentative.

Light therapy studies have consistently evaluated treatment response in terms of statistically significant reductions between pre- and posttreatment Hamilton Depression Rating Scale (HAM-D) scores, some allowing reduction of as little as four points as an indication of response in individual patients (e.g., Rosenthal et al. 1985a). Such demonstration of pre- to posttreatment change is dubious for demonstrations of efficacy because significant score reduction is common under placebo treatment (Klein et al. 1980; Fairchild et al. 1986) and in medicated patients judged clinically to be nonresponders.

The present cross-center analysis of treatment efficacy considers all formal investigations conducted from the start of this work in the early 1980s through the winter 1986–1987 season, involving 332 patients (Table 1). Contributing centers were located in 11 communities located at 39°N latitude or above throughout the United States, England, and Switzerland. Sample sizes of these studies, many as yet unpublished, have varied between 6 and 25 (median $n = 10$), precluding firm conclusions in many cases. Similarity of screening and assessment criteria, and of light therapy protocols, however, permits pooling of the present data, vastly increasing the sample size and moving toward more statistically confident comparisons of various light therapy regimens. To our knowledge, the data are comprehensive, with no controlled studies prior to 1988 excluded. We apply a set of definitions of treatment efficacy, varying in stringency, to pooled data of individual subjects under similar light exposure regimens across studies. The analysis is based on the "clustering approach" (Cochran 1963; Light and Smith 1971), in which independent subsamples are made up of subjects from separate studies who undergo a common manipulation. The resulting larger samples allow a statistically more powerful view of the time-of-day effect under morning, midday, evening, and morning-plus-evening light. These results were compared with parallel analyses of two control procedures (dim light and briefly presented bright light) in order to evaluate rel-

ative efficacy of presumed active and inactive treatments.

## METHODS

### Data Survey

We chose 21-item HAM-D scores, and measures derived from them, as dependent variables for analysis because the HAM-D is the only validated scale that has been consistently used. However, the scale has the drawback of underestimating clinical severity when used without addenda for the atypical vegetative symptoms of SAD, as incorporated into the SIGH-SAD (Williams et al. 1988). HAM-D scores were culled from published tables or figures; unpublished data were supplied by the investigators.

All studies used inclusion criteria for SAD patients similar to those of Rosenthal et al. (1985a): a history of major affective disorder as specified by the Research Diagnostic Criteria (Spitzer et al. 1978); at least two consecutive years of winter onset and springtime remission; absence of other (DSM-III Axis I) psychiatric disorders; and absence of seasonally recurrent psychosocial precipitants. Initial admission into the studies generally required a score of 14 or higher on the HAM-D scale, a value that might change at subsequent pretreatment baseline assessment. Although supplementary scales for atypical symptom severity were often administered (assessing fatigability, social withdrawal, appetite increase, increased eating, carbohydrate craving, weight gain, and hypersomnia; cf. Jacobsen and Rosenthal 1986), only Wirz-Justice et al. (1986) specifically required at least three such symptoms for inclusion. At present, investigators are adopting modified inclusion criteria for subsequent studies based on DSM-III-R (American Psychiatric Association 1987) categorization of mood syndromes and diagnosis of seasonal pattern. Because identification of psychosocial precipitants of winter depression is problematic, this exclusion criterion is being discontinued. Several research groups have excluded subjects using antidepressant medications, eliminating a potential confound with light therapy [though Wirz-Justice (personal communication) has noted no interaction with tricyclics]. As a precaution against possible adverse effects of bright light exposure on the eyes or skin, our group has also excluded subjects with histories of retinal disease, glaucoma, cataracts, diabetes, or skin cancer.

Because duration of light exposure during treatment differed among studies, separate analyses were performed for those reporting positive treatment effects with 2 to 6 hours of total daily exposure, and those using 1 hour of exposure or less, as a control or "low-dose" manipulation. The analysis focused on baseline and posttreatment comparisons, without consideration of withdrawal effects, due to dissimilar or absent washout procedures among the protocols.

### Data Analysis

***Measures of Improvement.*** Five measures, derived from the pre- and posttreatment HAM-D scores, were used to compare results within lighting conditions, by study, and for pooled data across conditions:

1. *Effect size* (*d* and *h*; Cohen 1977)—inferential statistics used in interpreting significance levels, and reported without sign. Effect size measures the influence of the treatment applied, taking into account the variability of the sample, and can be thought of as the difference between means in units of standard deviation. It can be used either to compare outcome scores between an "effective" treatment and a control or, as in Table 2, as a general indicator of the influence of a treatment on baseline scores. *d* was calculated from means and pooled standard deviations of raw scores using a computer program written by C. Patrick and J. Cohen (personal communication). *h* was taken from tables for proportional outcomes (Cohen 1977).
2. *Treatment-to-baseline ratio* ("post/pre" in Table 2)—provides a continuous scale of relative improvement that compensates for differences in subjects' baseline scores (Hamilton 1982).
3. *Relative improvement criterion*—the proportion of subjects showing a baseline-to-treatment reduction in HAM-D score of at least 50%.
4. *Absolute improvement criterion*—posttreatment HAM-D score less than 8. On a fundamental level, a treatment can be judged effective to the extent that the patient becomes symptom-free, regardless of relative improvement. It is our experience that patients with minimal residual symptomatology generally have scores below 8 on the HAM-D scale.
5. *Joint criteria (relative and absolute)*—applying these relative and absolute criteria, a strict and reliable gauge of symptom remission is obtained, permitting optimum discrimination among treatment outcomes. A subject who shows a 50% reduction in pre- to posttreatment HAM-D score, along with a posttreatment score under 8, can be considered to have shown sufficient magnitude of improvement not to require supplementary treatment. Ideally, this conclu-

**Table 1.** Studies Included in Cross-Center Analysis

| Source | Location | Total $n^a$ | Light exposure regimen | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Morning alone | Evening alone | Morning plus evening | Midday alone | Dim control | Brief control |
| D. Avery (p.c.)[b] | Seattle, WA | 7 | 7 | 7 | | | | |
| Checkley et al. 1986 | London, England | 11 | | | 11 | | 11 | |
| R. Depue (p.c.) | Minneapolis, MN | 15 | | | 15 | | 15 | |
| J. Docherty (p.c.) | Nashua, NH | 8 | 8 | | | | | |
| K. Doghramji (p.c.) | Philadelphia, PA | 6 | | 6 | | | | |
| C. Eastman (p.c.) | Chicago, IL | 6 | 6 | 6 | | | | |
| Hellekson et al., 1986 | Fairbanks, AK | 6 | 6 | 6 | 6 | | | |
| Hellekson and Rosenthal 1986 | | 7 | 7 | 7 | | 7 | | |
| C. Hellekson (p.c.) | | 7 | 7 | | | | | 7 |
| Isaacs et al. 1986 | London, England | 11 | | | 11 | 11 | 11 | |
| Jacobsen et al. 1987 | Bethesda, MD | 16 | 16 | | | 16 | | |
| James et al. 1985 | Bethesda, MD | 9 | | 9 | | | 9 | |
| Lewy et al. 1987 | Portland, OR | 8 | 8 | 8 | 8 | | | |
| A. Lewy (p.c.) | | 6 | 6 | 6 | | | | |
| R. McGrath (p.c.) | Teaneck, NJ | 11 | | 11 | | | | |
| Rosenthal et al. 1984 | Bethesda, MD | 11 | | | 11 | | 9 | |
| Rosenthal et al. 1985 | | 17 | | 11 | 17 | | 16 | |
| Rosenthal et al. (p.c.) | | 16 | | | 16 | | | |
| Sack et al. 1987 | Portland, OR | 10 | 10 | 10 | | | | |
| Sack et al. (p.c.) | | 14 | 14 | | | | | 14 |
| Terman et al. 1987 | New York, NY | 13 | 13 | 10 | | | | |
| Terman et al. 1989b | | 25 | 17 | 12 | 25 | | | 21 |
| Wehr et al. 1986a | Bethesda, MD | 7 | | | 7 | | | |
| Wehr et al. 1986b | | 10 | | 10 | | | | |
| Wirz-Justice et al. 1986 | Switzerland | 9 | | | 9 | | 6 | |
| Wirz-Justice et al. 1987 | | 25 | 25 | | | | | 23 |
| Wirz-Justice et al. (p.c.) | | 17[c] | 8 | 9 | | | | |
| Yerevanian et al. 1986 | Rochester, NY | 6 | 6 | 5 | | | | |
| Yerevanian et al. 1987 | | 18[c] | 8 | 10 | | | | |
| Totals | | 332 | 172 | 143 | 136 | 34 | 77 | 65 |

[a] Number of subjects whose data were considered in the cross-center analysis. *n* given in publications by the investigators sometimes differs due to their inclusion of partial data inadequate for present purposes. In the absence of documentation, we have assumed independent sampling across studies within a research center and believe there are only a few cases in which a subject participated in more than one experiment or was previously treated openly with light.

[b] p.c., personal communication. Several of the unpublished studies with $n < 10$ are accumulating additional subjects as of this writing, and the investigators do not consider their interim results to be conclusive.

[c] Separate groups design. All other studies with multiple conditions used complete or partial crossovers; in several cases, subjects were crossed over into conditions not considered in the present analysis.

sion would be supported by consideration of atypical vegetative symptoms of SAD, and Clinical Global Impressions. It is our impression that patients meeting the joint criteria usually exhibit minimal atypical posttreatment symptomatology as well.

## Statistical Tests

1. *The Wilcoxon signed-ranks test for matched pairs* (two-tailed; Siegel 1956) was used to determine significance of difference between pre- and posttreatment HAM-D scores in each lighting condi-

**Table 2.** Treatment Outcome Across Light Therapy Studies[a]

| Study | *n* | Light exposure schedule | HAM-D scale[b] Pre-treatment | HAM-D scale[b] Post-treatment | Wilcoxon *T*[c] | Effect size (*d*) | Ratio post/pre[d] | Proportion showing HAM-D decrements posttreatment −50% | <8 | −50%, <8 |
|---|---|---|---|---|---|---|---|---|---|---|
| **Morning Light** | | | | | | | | | | |
| D. Avery (p.c.) | 7 | 0600–0800 | 18.3 | 5.0 | * | 2.80 | 0.32 | 0.71 | 0.71 | 0.71 |
| J. Docherty (p.c.) | 8 | 0600–0800 | 18.1 | 6.5 | * | 1.88 | 0.34 | 0.75 | 0.75 | 0.75 |
| C. Eastman (p.c.) | 6 | 2 hr morning | 12.8 | 11.3 | NS | 0.37 | 0.92 | 0.17 | 0.17 | 0.17 |
| Hellekson et al. 1986 | 6 | 2 hr morning | 18.8 | 5.2 | * | 2.67 | 0.34 | 0.83 | 0.67 | 0.67 |
| Hellekson and Rosenthal 1986 | 7 | 2 hr morning | 19.1 | 10.0 | * | 1.43 | 0.62 | 0.57 | 0.41 | 0.41 |
| C. Hellekson (p.c.) | 7 | 2 hr morning | 25.1 | 14.7 | NS | 1.21 | 0.74 | 0.57 | 0.29 | 0.29 |
| Jacobsen et al. 1987 | 16 | 2 hr morning | 22.4 | 16.3 | ** | 0.98 | 0.75 | 0.31 | 0 | 0 |
| Lewy et al. 1987 | 8 | 0600–0800 | 15.4 | 6.6 | * | 1.50 | 0.42 | 0.63 | 0.63 | 0.63 |
| A. Lewy (p.c.) | 6 | 0600–0800 | 10.3 | 1.8 | * | 2.09 | 0.23 | 0.83 | 1.00 | 0.83 |
| Sack et al. 1987 | 10 | 0600–0800 | 11.8 | 4.3 | ** | 2.22 | 0.44 | 0.70 | 0.80 | 0.70 |
| R. Sack (p.c.) | 14 | 0600–0800 | 18.3 | 6.6 | ** | 1.75 | 0.40 | 0.86 | 0.50 | 0.50 |
| Terman et al. 1987 | 13 | 0600–0800 | 17.8 | 4.8 | ** | 2.81 | 0.28 | 0.77 | 0.77 | 0.77 |
| Terman et al. 1989b | 17 | 0600–0800 | 15.4 | 5.9 | *** | 2.06 | 0.36 | 0.71 | 0.76 | 0.71 |
| Wirz-Justice et al. 1987 | 25 | 0600–0800 | 20.4 | 10.4 | *** | 1.47 | 0.51 | 0.64 | 0.48 | 0.48 |
| A. Wirz-Justice (p.c.) | 8 | 0600–0800 | 17.5 | 6.9 | * | 2.54 | 0.42 | 0.75 | 0.50 | 0.50 |
| Yerevanian et al. 1986 | 6 | 0530–0730 | 16.7 | 4.3 | * | 2.53 | 0.31 | 0.83 | 0.83 | 0.83 |
| Yerevanian et al. 1987 | 8 | 0530–0730 | 15.4 | 8.6 | * | 0.94 | 0.52 | 0.50 | 0.63 | 0.50 |
| **POOL**[e] | **172** | | **17.8** | **8.1** | **** | **1.48** | **0.46** | **0.66** | **0.56** | **0.53** |
| **Midday Light** | | | | | | | | | | |
| Hellekson and Rosenthal 1986 | 7 | 2 hr midday | 20.3 | 7.3 | * | 2.63 | 0.39 | 0.71 | 0.57 | 0.57 |
| Isaacs et al. 1988 | 11 | 1000–1400 | 16.2 | 7.6 | ** | 1.37 | 0.48 | 0.55 | 0.45 | 0.45 |
| Jacobsen et al. 1987 | 16 | 1200–1400 | 21.6 | 14.6 | ** | 0.98 | 0.70 | 0.31 | 0.13 | 0.13 |
| **POOL**[e] | **34** | | **21.2** | **12.4** | **** | **1.25** | **0.56** | **0.50** | **0.32** | **0.32** |
| **Evening Light** | | | | | | | | | | |
| D. Avery (p.c.) | 7 | 2000–2200 | 19.4 | 15.1 | NS | 0.84 | 0.86 | 0.14 | 0.14 | 0.14 |
| K. Doghramji (p.c.)[f] | 6 | 1800–2000 | 21.0 | 9.8 | * | 2.01 | 0.50 | 0.50 | 0.50 | 0.50 |
| | 6 | 1800–2200 | 20.3 | 9.3 | * | 2.20 | 0.48 | 0.33 | 0.33 | 0.33 |
| C. Eastman (p.c.) | 6 | 2 hr evening | 12.8 | 17.0 | NS | 1.31 | 1.39 | 0 | 0 | 0 |
| Hellekson et al. 1986 | 6 | 2 hr evening | 17.5 | 4.7 | * | 2.38 | 0.30 | 0.83 | 0.83 | 0.67 |
| Hellekson and Rosenthal 1986 | 7 | 2 hr evening | 21.1 | 12.3 | * | 1.71 | 0.62 | 0.57 | 0.14 | 0.14 |
| James et al. 1985 | 9 | 1800–2300 or 1900–2400 | 24.3 | 13.4 | * | 1.26 | 0.56 | 0.44 | 0.33 | 0.33 |
| Lewy et al. 1987 | 8 | 2000–2200 | 15.4 | 15.3 | NS | 0.02 | 1.18 | 0.13 | 0.13 | 0.13 |
| A. Lewy (p.c.) | 6 | 2000–2200 | 10.3 | 8.2 | NS | 0.41 | 1.16 | 0.50 | 0.50 | 0.33 |
| R. McGrath (p.c.) | 11 | 2 hr evening | 15.4 | 7.4 | * | 1.35 | 0.57 | 0.55 | 0.45 | 0.36 |
| Rosenthal et al. 1985 | 11 | 1730–2030 | 19.6 | 10.8 | ** | 1.26 | 0.61 | 0.45 | 0.45 | 0.27 |
| Sack et al. 1987 | 10 | 1900–2100 | 11.8 | 7.9 | * | 0.83 | 0.75 | 0.30 | 0.50 | 0.30 |
| Terman et al. 1987 | 10 | 1900–2100 | 17.3 | 6.4 | ** | 2.38 | 0.36 | 0.80 | 0.70 | 0.70 |
| Terman et al. 1989b | 12 | 1800–2000 | 16.8 | 13.5 | NS | 0.64 | 0.84 | 0.17 | 0.17 | 0.17 |
| Wehr et al. 1986b | 10 | 1830–2230 | 25.6 | 10.3 | ** | 1.93 | 0.41 | 0.70 | 0.50 | 0.50 |
| A. Wirz-Justice (p.c.) | 9 | 2100–2300 | 16.0 | 6.7 | ** | 3.49 | 0.41 | 0.78 | 0.67 | 0.67 |
| Yerevanian et al. 1986 | 5 | 2000–2200 | 16.4 | 7.4 | NS | 1.10 | 0.39 | 0.80 | 0.80 | 0.80 |
| Yerevanian et al. 1987 | 10 | 2000–2200 | 18.4 | 6.7 | ** | 2.44 | 0.37 | 0.80 | 0.70 | 0.70 |
| **POOL**[e] | **143** | | **18.0** | **10.1** | **** | **1.16** | **0.62** | **0.50** | **0.43** | **0.38** |

(*continued*)

**Table 2.** *(continued)*

| Study | *n* | Light exposure schedule | HAM-D scale[b] Pre-treatment | HAM-D scale[b] Post-treatment | Wilcoxon *T*[c] | Effect size (*d*) | Ratio post/pre[d] | Proportion showing HAM-D decrements posttreatment −50% | <8 | −50%, <8 |
|---|---|---|---|---|---|---|---|---|---|---|
| **Morning Plus Evening Light** | | | | | | | | | | |
| Checkley et al. 1986[g] | 11 | 0700–1000 2000–2300 | 29.5 | 15.1 | ** | 1.18 | 0.46 | 0.55 | 0.45 | 0.45 |
| | 10 | 0700–0800 2200–2300 | 29.0 | 21.3 | ** | 0.79 | 0.73 | 0.20 | 0 | 0 |
| R. Depue (p.c.) | 15 | 0600–0800 1800–2000 | 25.6 | 6.5 | *** | 5.02 | 0.25 | 1.00 | 0.73 | 0.73 |
| Hellekson et al. 1986 | 6 | 1 hr morning 1 hr evening | 16.8 | 7.0 | ** | 2.92 | 0.44 | 0.83 | 0.67 | 0.67 |
| Isaacs et al. 1986 | 11 | 0500–0700 1700–1900 | 13.9 | 9.5 | * | 0.69 | 0.62 | 0.45 | 0.55 | 0.45 |
| Lewy et al. 1987 | 8 | 0600–0800 2000–2200 | 15.4 | 8.6 | NS | 0.96 | 0.71 | 0.63 | 0.50 | 0.50 |
| Rosenthal et al. 1984 | 11 | 3 hr morning 3 hr evening | 18.6 | 7.7 | ** | 2.15 | 0.40 | 0.64 | 0.54 | 0.54 |
| Rosenthal et al. 1985 | 17 | 0500–0800 1730–2030 | 24.4 | 12.8 | *** | 1.51 | 0.54 | 0.53 | 0.24 | 0.24 |
| Rosenthal et al. (p.c.) | 16 | 2 hr morning 2 hr evening | 18.2 | 10.1 | ** | 1.58 | 0.58 | 0.56 | 0.44 | 0.44 |
| Terman et al. 1989b | 25 | 0600–0800 1800–2000 | 16.2 | 4.7 | **** | 2.87 | 0.30 | 0.80 | 0.80 | 0.80 |
| Wehr et al. 1986a[h] | 7 | 0730–1030 2000–2300 | 29.1 | 14.7 | * | 1.63 | 0.62 | 0.71 | 0.14 | 0.14 |
| | 7 | 0900–1200 1400–1700 | 26.0 | 15.7 | * | 1.55 | 0.62 | 0.43 | 0.29 | 0.29 |
| Wirz-Justice et al. 1986 | 9 | 0600–0800 1800–2000 | 30.0 | 10.1 | ** | 2.70 | 0.39 | 0.67 | 0.33 | 0.33 |
| **POOL[e]** | **136** | | **21.1** | **9.2** | **** | **1.53** | **0.46** | **0.68** | **0.52** | **0.51** |
| **Dim Light Control[i]** | | | | | | | | | | |
| Checkley et al. 1986 | 11 | 0700–1000 2000–2300 | 27.4 | 23.2 | NS | 0.44 | 0.82 | 0.27 | 0 | 0 |
| R. Depue (p.c.) | 15 | 0600–0800 1800–2000 | 25.6 | 23.9 | * | 0.39 | 0.94 | 0 | 0 | 0 |
| Isaacs et al. 1986 | 11 | 0500–0700 1700–1900 | 16.3 | 11.7 | NS | 0.76 | 0.75 | 0.27 | 0.27 | 0.18 |
| James et al. 1985 | 9 | 1800–2300 or 1900–2400 | 23.6 | 18.6 | NS | 0.60 | 0.78 | 0.22 | 0.22 | 0.22 |
| Rosenthal et al. 1984 | 9 | 3 hr morning 3 hr evening | 15.1 | 13.2 | NS | 0.32 | 0.86 | 0.33 | 0.22 | 0.22 |
| Rosenthal et al. 1985[j] | 10 | 0500–0800 1730–2030 | 20.3 | 18.1 | NS | 0.32 | 0.91 | 0.20 | 0.20 | 0.20 |
| | 6 | | 25.3 | 26.5 | NS | 0.18 | 1.11 | 0 | 0 | 0 |
| Wirz-Justice et al. 1986 | 6 | 0600–0800 1800–2000 | 26.3 | 13.5 | * | 1.72 | 0.54 | 0.50 | 0.16 | 0.16 |
| **POOL[e]** | **77** | | **23.4** | **20.0** | *** | **0.44** | **0.85** | **0.21** | **0.13** | **0.11** |
| **Brief Exposure Control—Bright Light[k]** | | | | | | | | | | |
| C. Hellekson (p.c.) | 7 | 0.5 hr morning | 27.9 | 23.6 | NS | 0.55 | 0.83 | 0.14 | 0 | 0 |
| R. Sack (p.c.) | 14 | 0600–0630 | 18.3 | 8.4 | ** | 1.62 | 0.49 | 0.57 | 0.50 | 0.50 |
| Terman et al. 1989b | 21 | 0730–0800 1800–1830 | 15.5 | 8.7 | *** | 1.48 | 0.58 | 0.33 | 0.29 | 0.29 |
| Wirz-Justice et al. 1987 | 23 | 0600–0630 | 19.9 | 11.9 | *** | 1.30 | 0.59 | 0.43 | 0.30 | 0.30 |
| **POOL[e]** | **65** | | **19.0** | **11.4** | **** | **1.08** | **0.59** | **0.40** | **0.32** | **0.31** |

*(continued)*

**Table 2.** (*continued*)
NOTES.

[a] All bright-light studies used full-spectrum fluorescent lamps (Vita-Lite) providing approximately 2500 lux at the level of the eyes, except for the studies of Yerevanian et al. (1986, 1987), which used incandescent lamps, and Lewy et al. (1987) in which some patients were given cool-white fluorescent lamps.

[b] Mean 21-item HAM-D scores are based on computations of individual data shown in published tables or graphics, or from unpublished original data provided by the investigators. [Three studies (Isaacs et al. 1986; Yerevanian et al. 1986; C. Eastman, personal communication) used the 17-item HAM-D scale. Yerevanian et al. (1987) used a 23-item version that included points for atypical vegetative symptoms. These data are omitted from tabulations of mean 21-item HAM-D scale scores in the cross-study pools. They are included, however, in pooled analyses of proportional improvement.] Whenever possible within a crossover design, withdrawal scores immediately preceding a light treatment condition are used as baseline; given successive treatments without withdrawals, the pretreatment score is used. Pooled data for each condition are based on individual subjects' scores, not on mean scores per study; studies with larger $n$ therefore contribute more heavily to the pooled results.

[c] Wilcoxon signed rank test for correlated samples (two-tailed): $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, $^{****}p < 0.0001$.

[d] Treatment-to-baseline ratios are based on individual subjects' matched scores.

[e] Pooled $n$ and pooled proportional measures exclude multiple entries for patients who were tested in more than one group at a given time of day. Pooled means and Wilcoxon tests exclude such multiple entries, as well as studies that did not use the 21-item HAM-D scale (see note *b*).

[f] Patients tested in a crossover between 2 and 4 hr evening light. Only the 4-hr data are entered into the pooled analyses.

[g] Patients tested in a crossover between 1 and 3 hr morning-plus-evening light. Only the 3-hr data are entered into the pooled analyses, given minimal evidence for efficacy at 1 hr among this hospitalized patient group.

[h] Patients tested in a crossover between "long" and "short" photoperiod conditions. Only the long photoperiod data are entered into the pooled analyses; these correspond most closely with the morning and evening exposure intervals of the other experiments.

[i] Dim light parameters: Checkley et al. (1986), 400 lux red incandescent; Depue (personal communication), 300 lux full-spectrum fluorescent; Isaacs et al. (1988), 300 lux incandescent; James et al. (1985), 300 lux full-spectrum fluorescent; Rosenthal et al. (1984), 100 lux yellow fluorescent; Rosenthal et al. (1985), 300 lux full-spectrum fluorescent (ten outpatients), 5 lux full-spectrum fluorescent (six inpatients); Wirz-Justice et al. (1986), 250 lux yellow fluorescent.

[j] This study included subgroups of ten outpatients and six inpatients (results reported separately).

[k] Thirty-minute sessions using 2500 lux full-spectrum fluorescent light.

tion, by study. This test was chosen rather than the more powerful Student's $t$ test because the nonnormality of several studies' score distributions violated the assumptions of parametric testing. This test was also used to assess significant movement between pre- and posttreatment HAM-D scores among subjects who failed to achieve clinical remissions in morning–evening light crossover studies.

2. *The Kruskal–Wallis one-way analysis of variance* (Siegel 1956) was used to test significance of difference between the six pooled group lighting conditions (morning, midday, evening, morning plus evening, dim, and brief) for, separately, baseline, posttreatment, and relative improvement ("post/pre") scores. This nonparametric procedure was chosen because significant outcomes of the Bartlett–Box test for homogeneity of variances between the groups (Bock, 1975) violated, for both baseline and posttreatment scores, assumptions of the parametric $F$ test. As a rule of thumb, in a situation of $J$ groups with unequal variances and unequal sample sizes, if $n$ and variances are substantially negatively correlated, use of a parametric test will result in excessive Type 1 error (J. Cohen, personal communication). The correlation in the case of the pooled group posttreatment scores was negative and large ($r = -0.68$, NS; significance test limited by $df = 5$). The necessary inclusion of subjects who underwent more than one treatment condition results in inflated error variance, which acts to make this a conservative test of significance.

3. *The Mann–Whitney U test* (Siegel 1956) was used for significant differences in, separately, baseline, posttreatment, and relative improvement scores between any pairs of pooled group lighting conditions, with application of the Bonferroni inequality correction (Wilcox 1987) for multiple comparisons.

4. *McNemar's test for significance of changes* (Siegel 1956) was used among 71 subjects in six studies receiving both morning light alone and evening light alone in crossover designs, and for 66 subjects receiving bright and dim light under the morning-plus-evening schedule. Given limited power efficiency of this test, we include effect size ($h$) outcomes to assist in interpretation.

5. *The binomial test* (Siegel 1956) was used for significance of changes within crossover procedures lacking expected frequencies large enough to perform McNemar's test. As with McNemar's test, we include the $h$ statistic.

6. *The hypothesis test of differences in proportions* (Ferguson 1959) was used for subjects who met the various criteria for clinical improvement and remission under different lighting regimens. This test was also used to compare results for groups that differed in baseline symptom severity.

7. *McNemar's test of symmetry* (Bowker 1948) was

used for comparisons of posttreatment HAM-D score distribution shapes under morning and evening light in crossover studies.

8. *The Kolmogorov–Smirnov one-sample test* (Siegel 1956) was used for significant differences from normality in distribution shapes for morning and evening posttreatment scores in crossover studies.

## RESULTS

### Statistical Significance and Effect Size of Posttreatment Score Reductions

Table 2 presents treatment outcome measures across light therapy studies for the following lighting conditions: morning light ($n$ = 172); midday light ($n$ = 34); evening light ($n$ = 143); morning-plus-evening light ($n$ = 136); and two controls, dim light ($n$ = 77) and briefly presented bright light ($n$ = 65). The pooled mean HAM-D baseline scores ranged between approximately 18 and 23, and dropped to approximately 8 to 12 under the bright light treatments. These score reductions were highly statistically significant ($p < 0.0001$). Note, however, that only two of the conditions, morning light alone and morning-plus-evening light, yielded mean posttreatment scores under 10. Although the response to dim light showed a mean score reduction of only approximately 3 points to a posttreatment level of 20.0, this was also a highly statistically significant difference from baseline ($p < 0.001$). Thus, by the standard of statistical reduction of scores alone, all the treatment schedules would be considered effective.

The majority of individual studies comprising the clusters for all bright light treatments achieved significant score reductions, though six of the evening light studies showed no such difference, in comparison with only two under morning light. For dim light, only two of eight showed significant score reductions under treatment, despite the strong result using the larger, pooled sample size.

The usefulness of significance levels of score reductions between baseline and posttreatment scores is limited when considering studies with small sample sizes—there is insufficient power in a test of difference between populations to reject the null hypothesis even when the effect associated with the alternative hypothesis is substantial. Consideration of significance levels along with a measure of effect size, such as the $d$ statistic (Cohen 1977), gives a truer picture of outcome and helps avoid mistaken conclusions about relative efficacy.

Across the lighting regimens, effect sizes of individual studies show a large range reflecting variability of both score reduction and spread of scores. Effect size outcomes for the pooled data show that bright light conditions with morning light lead to mean posttreatment score reductions of approximately 1.5 SD units. Midday and evening light show somewhat smaller effect size (1.25 and 1.16 SD units, respectively), followed by the brief exposure control (1.08 SD units). The dim light control lowered scores less than 0.5 SD unit. In comparing morning and evening light, with evening as a control, the difference in effect is $d$ = 0.31. It should be noted that this normalized difference measure is compromised by pooling subjects from differing protocols, with a resultant increased score variability. When the morning–evening difference is examined with subjects receiving both conditions in crossover tests, the normalized difference is much larger (see results for matched-pairs subsets, below).

Mean baseline HAM-D scores for the pools did not differ statistically from one another in most cases, permitting direct comparisons of posttreatment scores without the confound of variable pretreatment severity. A major exception, however, was the dim light baseline (23.4), which was significantly higher than baselines of morning alone, evening alone, and the brief exposure control (Mann–Whitney U, $p < 0.05$ with Bonferroni correction). In addition, both morning alone and evening alone showed lower baseline scores than morning plus evening ($p < 0.05$). In comparisons of mean posttreatment pooled scores, all of the bright light treatments separately yielded significantly lower values than that of the dim light control ($p < 0.05$). In addition, morning light alone showed a significant advantage over evening light alone ($p < 0.05$). If, however, the magnitude of a treatment effect depends on baseline severity, as we demonstrate below, direct comparisons of posttreatment scores would become problematic.

### Ratio Improvement of Posttreatment Scores

Following Hamilton's (1982) suggestion, baseline score variations were factored out by computation of a proportional improvement measure for each individual subject, the treatment-to-baseline ratio ("post/pre" in Table 2), and means were derived for each experiment and the pooled data. In general, effective antidepressant treatments are expected to yield ratios of 0.5 or less. (Hamilton noted that patients are not fully satisfied with a treatment unless this ratio falls below 0.33.) Within each light exposure regimen, variations in this ratio are quite large across studies (e.g., 0.23 to 0.75 for morning light alone). Yet the pooled results sustain the impression of greater benefit for morning light alone (0.46) and morning-plus-

evening light (0.46) in comparison to midday light (0.56) and evening light (0.62). In Mann–Whitney tests, the "post/pre" ratio showed a significant advantage of morning over evening light, morning-plus-evening light over evening light, and all of the bright light conditions, separately, over the dim light control ($p < 0.05$ with Bonferroni correction).

### Proportion of Patients Meeting "Remission" Criteria

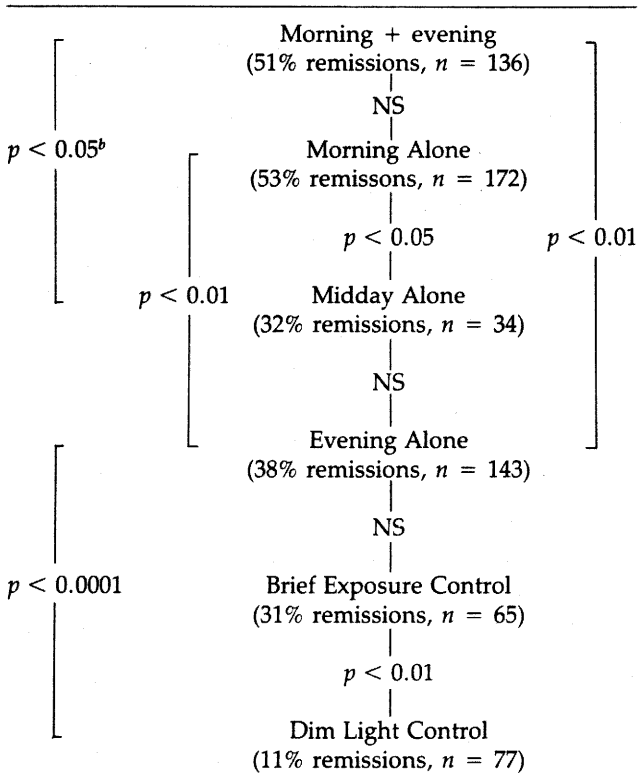
The data were further analyzed by setting three operational definitions of clinical remission, with increasing strictness, for determining the proportion of patients experiencing successful treatment under the various lighting regimens. The laxest criterion—the proportion showing at least 50% reduction in HAM-D score between baseline and posttreatment assessments—was derived from individual subjects' "post/pre" ratios. Next, an absolute criterion for remission was applied by determining the proportion achieving a posttreatment score under 8, considered to be within the range for normals. Finally, joint criteria were applied requiring that a patient show both 50% score reduction and a posttreatment score under 8 to be considered remitted. This most conservative measure rejects both severe cases, in which symptoms remained despite large proportional improvement (e.g., score reduction from 33 to 15), and mild cases, in which few symptoms remained after the therapy but in which proportional improvement was relatively small (e.g., score reduction from 12 to 7). In either situation, we argue, a strong case for antidepressant efficacy cannot be made.

***Relative Improvement Criterion.*** Across studies and across light exposure conditions the proportion of patients with HAM-D score reductions of 50% or greater shows a wide range (e.g., 0.17 to 0.86 within the 17 morning light studies). Small sample sizes are at least partly responsible for such wide variability. Yet orderliness is obtained by clustering. Morning-plus-evening light and morning light alone showed remission proportions of 0.68 and 0.66, respectively. Midday light alone and evening light alone both yielded 0.50, the brief exposure control 0.40, and the dim light control 0.21. Hypothesis tests for proportions showed statistically significant variations within this graded range. Prominently, morning light showed a significant advantage over both midday ($p < 0.02$) and evening ($p < 0.002$) light, which did not differ from each other, and bright light conditions were all more effective than the dim light control ($p < 0.0001$).

***Absolute Posttreatment Criterion.*** Mean pooled posttreatment HAM-D scores showed a fairly small range across the bright light treatments (morning light, 8.1; midday light, 12.4; evening light, 10.1; morning-plus-evening light, 9.2). All of these values border on the mildly symptomatic range, however, suggesting that half or more of any treatment group failed to achieve unambiguous remissions. The absolute remission criterion (posttreatment score under 8) was met by 0.56 of subjects under morning light alone, 0.52 under morning-plus-evening light, 0.32 under midday light, 0.43 under evening light, and 0.13 under dim light. Within this graded range, morning light alone did not differ significantly, in hypothesis tests of proportions, from morning-plus-evening light, and although morning light exceeded both midday ($p < 0.01$) and evening ($p < 0.05$) light, midday and evening light did not differ. Again, all bright light conditions, separately, exceeded the dim light control ($p < 0.0001$). It should be noted that the HAM-D score cutoff of 8, chosen for reasons of clinical utility, does not determine the effects noted here—when the cutoff was scaled upward, as high as 12, the results of these treatment comparisons were the same.

***Joint Criteria (Relative and Absolute).*** The estimated success rates are, of course, lowest, and most conservative, when both relative (50% reduction in HAM-D score) and absolute (HAM-D <8) remission criteria are jointly applied. A proportion of 0.53 of subjects who received morning light alone achieved remissions by this standard, with 0.51 under morning-plus-evening light. By comparison, the criteria were met by only 0.32 of subjects under midday light, and 0.38 of subjects under evening light. Table 3 summarizes the results of hypothesis tests for proportions across the treatment conditions. The two regimens with morning light exposures were equally effective and were superior to both midday and evening light. The most statistically significant difference was between bright light conditions, regardless of time of day, and the dim light control. With fewer than 40% remissions under midday and evening light (within the placebo range established for depressed patients in medication studies), an active antidepressant effect of light would seem questionable on average, whereas the morning procedures yield a moderately convincing result. The global results argue against an antagonistic effect of combined morning and evening exposures (cf. Lewy and Sack 1986); rather, when all studies are taken into account, the addition of evening light appears irrelevant to the treatment effect. The data clustering does not permit a contrast of early or late evening light, and phase delays in the latter case might prove selectively countertherapeutic. Results of individual evening light studies (Table 2), however, are not explained by the timing of light exposure in hours of dusk or later: The relatively effective evening treatments span the

**Table 3.** Proportion of "Complete Remissions"[a] Across Treatment Conditions




[a] Strict joint remission criteria: pre- to posttreatment reduction of Hamilton Depression Rating Scale score of at least 50% to a level under 8.

[b] Hypothesis test of differences in proportions.

same range of exposure times as those failing to demonstrate improvement.

Data of the individual studies largely corroborate the conclusions from the pooled results, with some exceptions. Of the 17 studies using morning light alone, 14 yielded a proportion above 0.40 of subjects meeting our joint criteria. One exception was the study of Jacobsen et al. (1987), in which no subject showed major improvement; the mean posttreatment score was 16.3, the highest of any morning-light study, a value that in fact exceeds the pretreatment baseline scores of six morning-light studies. The other exceptions (C. Eastman, personal communication, $n = 6$; C. Hellekson, personal communication, $n = 7$) were subject to small sample sizes. Of the 17 studies that presented evening light alone, 11 failed to produce a proportion of subjects responding greater than 0.40. The six studies that did exceed proportions above 0.40 included three with $n \leq 6$ (Hellekson et al. 1986; Yerevanian et al. 1986; K. Doghramji, personal communication).

### Subjects Within Crossover Designs

***Morning Versus Evening Light.*** One might wonder if the differential efficacy of morning versus evening light is due to protocol discrepancies across the individual studies. We therefore compared ten experiments from six centers that crossed subjects over between morning light alone and evening light alone ($n = 71$), permitting a matched-pairs analysis. The scatter plot in Figure 1a correlates individual posttreatment HAM-D scores. If morning light and evening light yielded equivalent effects, data points would cluster about the major diagonal. Instead, most points fall above this line, indicating generally higher posttreatment scores following evening light treatment.

The dashed lines in the plot demarcate the absolute criterion of HAM-D score under 8, separating scores considered to represent a symptomatology in the "normal" range from those above. Thirty-one percent of cases (22 of 71) meet this criterion, or remitted, after either morning or evening light treatment. Thirty-seven percent (26 of 71) show HAM-D scores under 8 after morning light, but have scores of 8 or higher after evening light. In contrast, only 6% (4 of 71) show a preferential response to evening light. Of this small group, only two of the patients, one from Oregon and one from Rochester, showed a strong evening-light bias and might be considered to belong to a small minority of "phase-advanced" SAD patients requiring evening light exposure for a normalizing delay (cf. Lewy et al. 1988). For the 27% of cases (19 of 71) with posttreatment HAM-D scores of 8 or higher under both treatments, the tendency toward stronger response to morning light is maintained, and it is apparent that the effect would be sustained if the absolute cutoff score of 8 were shifted upward.

The morning–evening difference was verified statistically for these matched pairs by comparing the proportion of subjects satisfying our joint criteria for remission—HAM-D score reduction of 50% or more to a level under 8—under morning light alone and evening light alone (McNemar's $\chi^2 = 18.89$, $df = 1$, $p < 0.001$). Overall, 62% (44 of 71) of these crossover patients showed remissions under morning light alone, in contrast to 28% (20 of 71) under evening light alone, magnifying the difference in proportions found in Table 2 for the total pool (53% and 38%, respectively). Fifty-nine percent (26 of 44) of morning light responders failed to respond to evening light. In comparison, only 10% (2 of 20) of evening-light responders failed to respond also to morning light. If, in future research, the therapeutic effect of light proves to be conditional on morning or evening exposure in phase-delayed and phase-advanced patients, respectively, nondifferential positive responders in morning–evening crossover studies, 25%, might be taken reasonably to be placebo responders.

When one considers the effect size measures for the matched pair subjects in comparison to those for

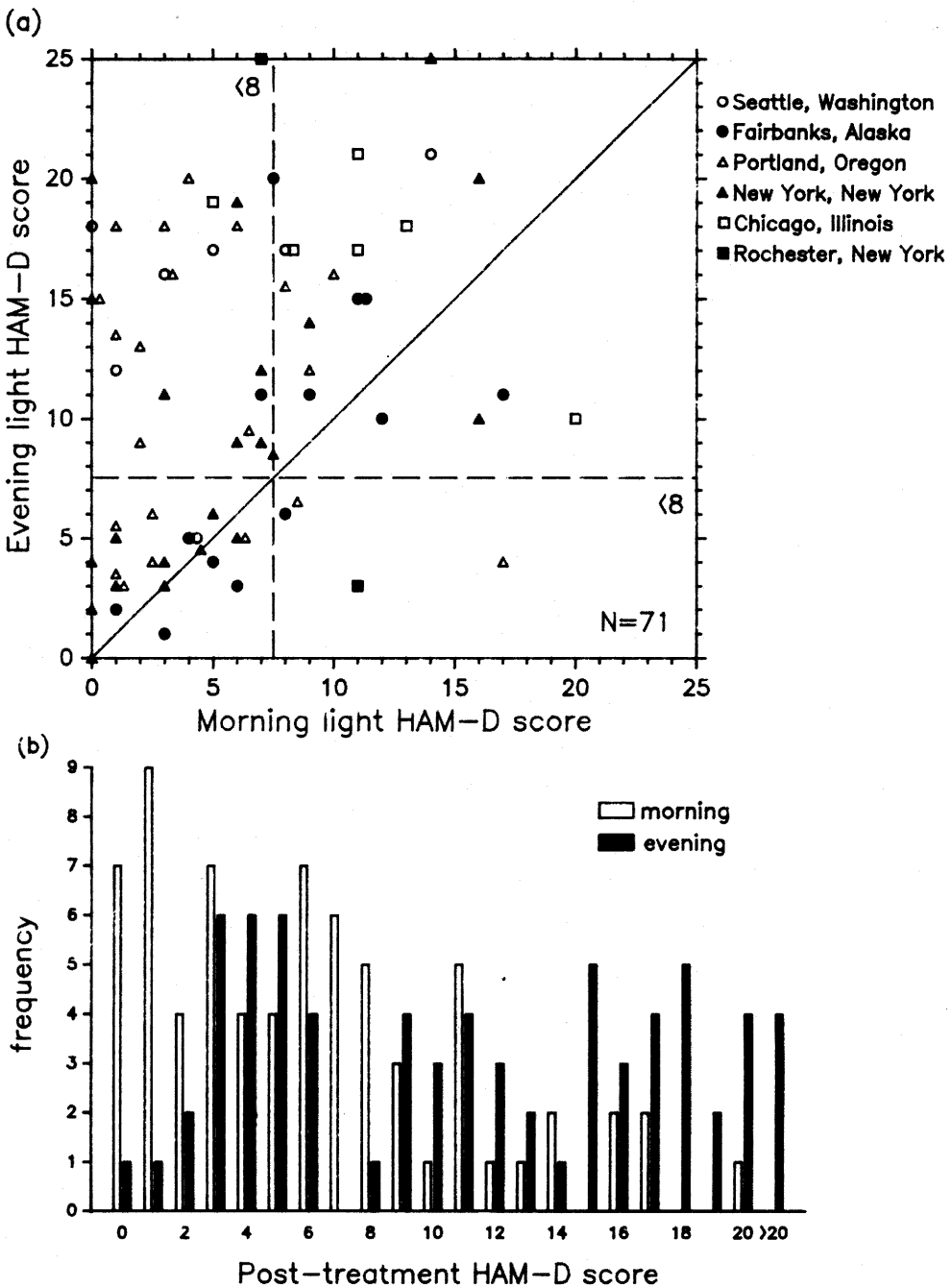



**Figure 1.** **(a)** Correlation of posttreatment Hamilton Depression Rating Scale (HAM-D) scores for individual patients at six research centers given morning light and evening light treatment in crossover protocols. Points falling above the major diagonal indicate a preponderance of differentially positive responses to morning light. Points falling below the criterion cutoff score of 8 (*dashed lines*) serve to estimate remission rate to morning light alone, evening light alone, or both (overlapping quadrant at lower left). **(b)** Frequency distributions of posttreatment HAM-D scores for subjects crossed over between morning and evening light regimens.

the total pool, the decrease in variability leads to an even more striking display of differential efficacy between morning and evening light: difference in means leads to a $d$ of 0.87. When scores are filtered by the joint criteria for complete remissions, the difference is still substantial ($h = 0.70$).

The shape of morning and evening posttreatment HAM-D score distributions (Fig. 1b) also differs, as shown by relative frequencies categorized according to severity level: low, 0–9; moderate, 10–16; and high, 17 or above (McNemar's test of symmetry, $p < 0.0001$). Furthermore, the evening

score distribution is not normal (Kolmogorov–Smirnov one sample goodness of fit test, $p < 0.05$). Evening scores are made up of two distinct subsets: one third of cases falling at the low end of the posttreatment severity range (HAM-D score under 8), and two thirds spread with essentially equal probability across the symptomatic range (HAM-D score 8 or higher). The fact that the latter subset shows no significant movement of HAM-D scores after treatment (Wilcoxon signed ranks test, NS) suggests that these patients were "complete nonresponders" to evening light. This stands in contrast with the morning-light distribution, with approximately two thirds of cases showing posttreatment HAM-D scores under 8, and where, among those who still fall within the symptomatic range there is, however, a significant decrease in HAM-D scores after treatment ($p < 0.01$). The improvement seen in morning-light nonresponders suggests dose dependency for morning light exposure: had this light been more intense or presented for sufficient duration, the morning nonresponders might have shown remissions.

Given between-subject variability in the data for all studies, in contrast to the crossovers, one should not test for corresponding effects using pooled raw scores. However, using the joint criteria for remission (pre- to posttreatment HAM-D score reduction of 50% or above to a level under 8), which help control for such variability, the distinction between morning and evening posttreatment distribution shapes is similar to that seen for the matched pairs. The evening light distribution shows a broad spread of proportions, with only 39% of studies resulting in a majority of patients meeting the joint criteria. This contrasts with a normal distribution of morning-light proportions wherein 71% of studies yield a majority showing remissions.

***Other Matched-Pairs Comparisons.*** By similar assessment of 28 patients who received both morning light alone and morning-plus-evening light in crossover tests (Table 1), there was no evidence of a differential effect ($h = 0.09$, binomial test for small samples, NS). Sixty-six subjects who were treated with both bright and dim light under the morning-plus-evening schedule showed a strong bright-light benefit ($h = 0.84$, $\chi^2 = 19.53$, $df = 1$, $p < 0.001$). In contrast, 18 subjects who were treated with both bright and dim light in the evening alone failed to show a statistically differentiated response ($h = 0.50$, binomial test, NS), albeit moderate effect size.

In general, these matched-pairs comparisons, although vulnerable to small sample sizes, sustain and amplify trends observed within the total pool (Table 2) by reducing subject variance inherent to the independent-groups analysis. The data indicate (1) the superiority of morning over evening light, (2) the lack of consequence—positive or negative—of adding evening to morning exposures, and (3) the superiority of bright light over the dim light control, given morning exposures, but no such clear-cut superiority given evening exposures.

## Treatment Response As a Function of Baseline Severity

The covariation of individual subjects' pre- and posttreatment scores is illustrated in the scatter plots of Figures 2a and 2b for the three presumed active treatment conditions that have been investigated most thoroughly—morning-plus-evening light ($n = 136$), morning light alone ($n = 172$), and evening light alone ($n = 143$)—and for the dim light control ($n = 77$). This display reveals considerable variation within and across light therapy studies, and systematic aggregation of points for a given research center or geographic area can be directly detected. The raw data display permits one to "cut the cake" by rules different from ours. Dividing lines demarcate our response criteria applied across the range of symptom severity.

The data are scattered widely throughout the space encompassing pre- to posttreatment score reductions. Under morning light alone the points appear to aggregate at relatively low pre- and posttreatment scores. This could result from a selection bias toward milder cases. The highest baseline scorers in evening-alone studies have been entirely from Bethesda, and the lowest scorers—several of whom showed symptom exacerbation under evening light—have been predominantly from Portland. In the main, however, there is a broad overlap of data points across the research centers, which supports the legitimacy of the clustering approach. The wide range of baseline HAM-D scores found within and across light therapy studies may cause concern that the subject population is nonhomogeneous. A high baseline score requires a larger response, both relatively and absolutely, for obtaining a remission. In order to ascertain whether our global conclusions about differential treatment efficacy describe the SAD population as a whole, we have further analyzed treatment outcome for morning-plus-evening light, morning light alone, evening light alone, and the dim light control with baseline symptom severity split into two groups: clinically mild cases (HAM-D score between 10 and 16) and moderate-to-severe cases (>16). Subjects with baseline scores under 10 were omitted from this analysis on the grounds that their low severity level would tend to confound assessment of antidepressant efficacy. (For the pooled

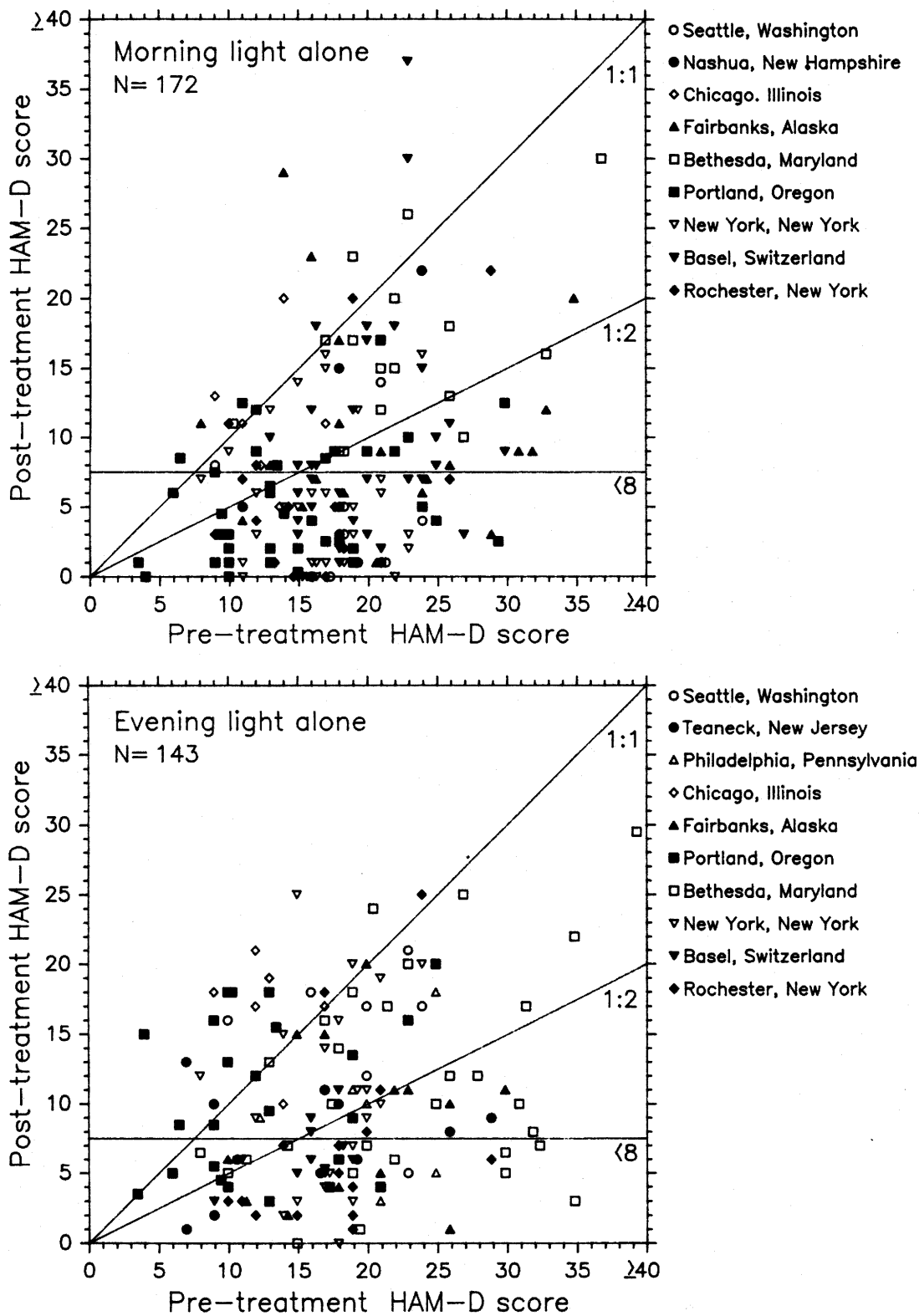



**Figure 2a.** Scatter plot of pre- and posttreatment HAM-D scores for individual patients treated with morning light alone and evening light alone. Points clustered about the major diagonal indicate no posttreatment reduction in symptom severity. Points near the 1:2 contour indicate relative posttreatment score reductions of approximately 50%. Points below 8 on the ordinate indicate posttreatment symptom severity within a normal, or subclinical range.

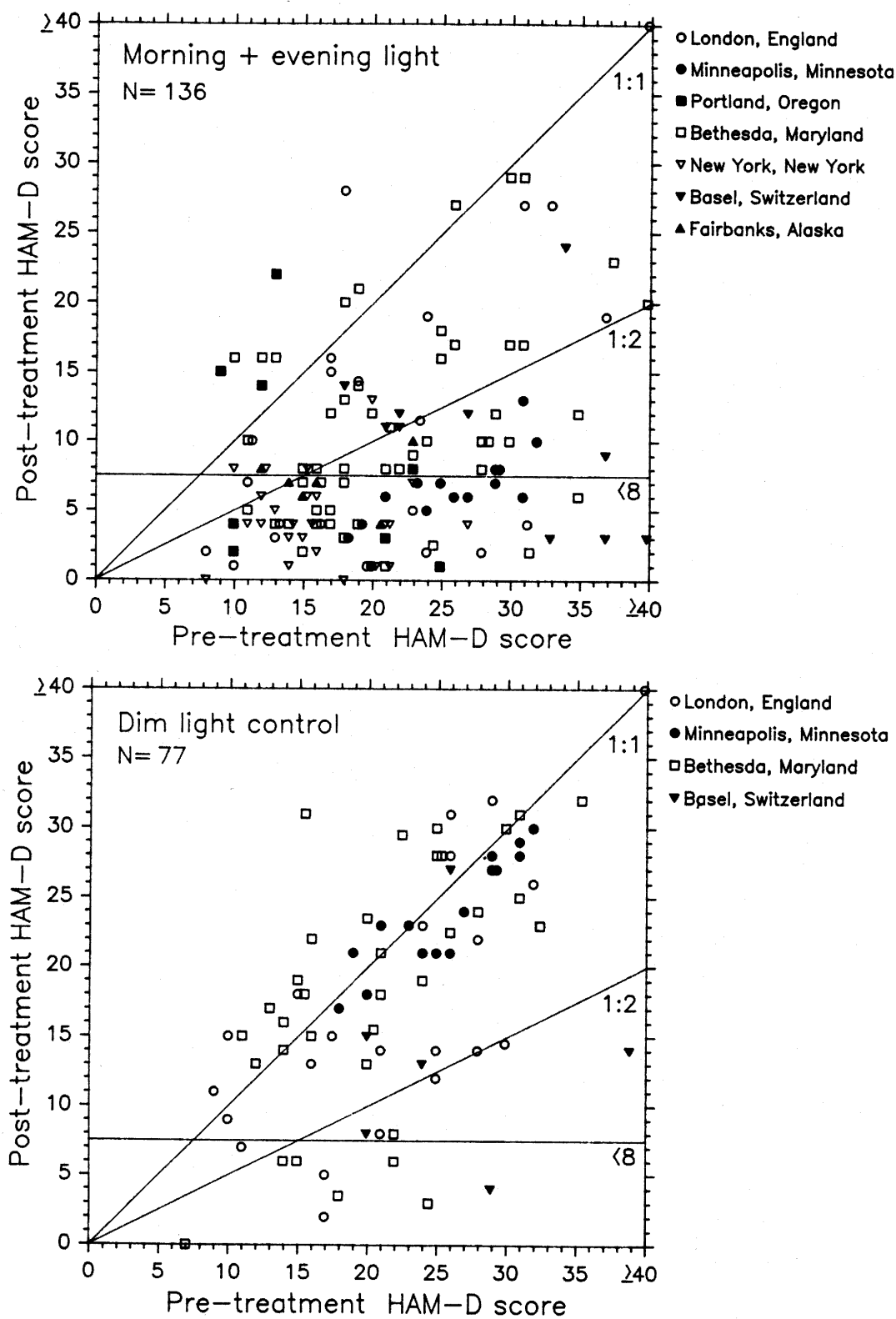



**Figure 2b.** Scatter plots, as in Figure 2a, of HAM-D scores for individual patients treated with the combination of morning-plus-evening light exposures, and a dim light control.

results in Table 2, however, results did not differ statistically whether or not such subjects were included. Despite higher baseline HAM-D requirements at initial entry, several studies found occasional substantial pretreatment score reductions, or failures to relapse during withdrawals separating crossovers, and yet proceeded with treatment, factors that could favorably bias the group-average response.)

Figure 3 plots the proportion of subjects meeting the joint criteria in four treatment conditions—morning-plus-evening, morning alone, evening alone, and the dim control—as a function of baseline severity. When mild and moderate-to-severe baseline ranges are pooled, evening light alone is observed to be significantly less effective than morning-plus-evening light and morning light alone (hypothesis tests of proportions, $p < 0.05$), and the dim light control is even less effective ($p < 0.0001$). When subjects in the mild baseline range are considered separately, the greater efficacy of morning over evening is considerably magnified, with 0.67 of subjects meeting the joint criteria under morning light alone, compared with 0.39 for evening light alone ($p < 0.01$). This morning–evening difference vanishes for the more severe baseline cases, and the efficacy of morning-plus-evening light and morning light alone fall significantly below that of the mild baseline cluster ($p < 0.05$ and $p < 0.01$, respectively). Thus, at the higher level of baseline severity, time of day of treatment has little effect on rate of remission. Both evening light and dim light response rates remain approximately equal (0.39 and 0.11, respectively) across baseline severity ranges. This analysis localizes the differential treatment effect of morning light to mild SAD cases. That the clinical effect of evening light and dim light does not vary with baseline severity further suggests that they are nonspecific or inactive treatments, despite the statistically significant reductions seen in the mean HAM-D scores (Table 2).

It might be argued that bisecting the level of baseline symptom severity at another HAM-D score cutoff would reveal an apparent baseline effect not seen with a cutoff of 16. When the analysis is applied to cutoffs up to 20, however, the lower efficacy of treatment in the more severe cases, and lack of significant difference in remission rates as a function of time of day, is maintained.

One might reasonably suspect that use of the joint remission criteria, which includes the requirement of posttreatment HAM-D score under 8, would bias conclusions in favor of mild cases in which baseline scores are already low. In contrast, use of a relative improvement criterion, such as the proportion of patients achieving a 50% reduction in HAM-D score, might obviate the baseline effect. We tested the differences between "mild" and "moderate-to-severe" subgroups for relative improvement and still found a significant advantage of morning over evening light given low baseline severity (0.67 vs. 0.43 proportions of patients with 50% score reduction, $p < 0.01$), an effect that decreased or vanished at moderate-to-high severity (0.66 vs. 0.57, NS).

Examination of Figure 1a discloses some evidence of this baseline effect for the matched-pairs data set, though limited by small sample sizes within studies and perhaps also (with the pooled mean HAM-D baseline scores under 16) an inclusion bias toward milder cases. Given undifferentiated efficacy for morning and evening light at moderate-to-high severity, shown in Figure 3, one would expect to find the more severe cases clustering around the diagonal and, given the marked difference at low severity, milder cases clustering to the left of the vertical cutoff line (posttreatment HAM-D score under 8). This seems to be the pattern for, most notably, the results of Fairbanks and Portland. The Alaskan HAM-D baselines, at 19 or above, are significantly higher than the Oregon baselines for both morning and evening (both $p < 0.05$, Mann–Whitney U test) and higher than the rest of the data set in the case of evening ($p < 0.05$). Considering then the Alaskan data ($n = 13$) to constitute a "high severity" sample, data points appear grouped relatively close to the diagonal and do not show a significant morning–evening difference in proportions meeting the joint criteria for remission, despite an apparent trend (0.58, morning; 0.38, evening). In contrast, the "low severity" Oregon subjects ($n = 21$, mean HAM-D baseline score = 13.2) showed significantly more remissions under morning than evening treatment (0.64 vs. 0.21, $p < 0.05$).

## DISCUSSION

The aggregate of analyses in this cross-center study show clearly that the clinical efficacy of light treatment varies significantly with time of day of exposure as well as intensity and baseline severity. Given the particulars of the most effective treatment combination—high-intensity light, early-morning exposure, mildly depressed cases—the argument for specific clinical efficacy is sustained.

Is evening light effective? Its average efficacy falls within the upper range of placebo response established for antidepressant medication studies, but those were not studies of SAD, so that a strict comparison would be false. Pooled crossover results for bright evening light vs. dim light controls (either evening alone or morning plus evening) show a moderate bright-evening differential effect. Several indi-

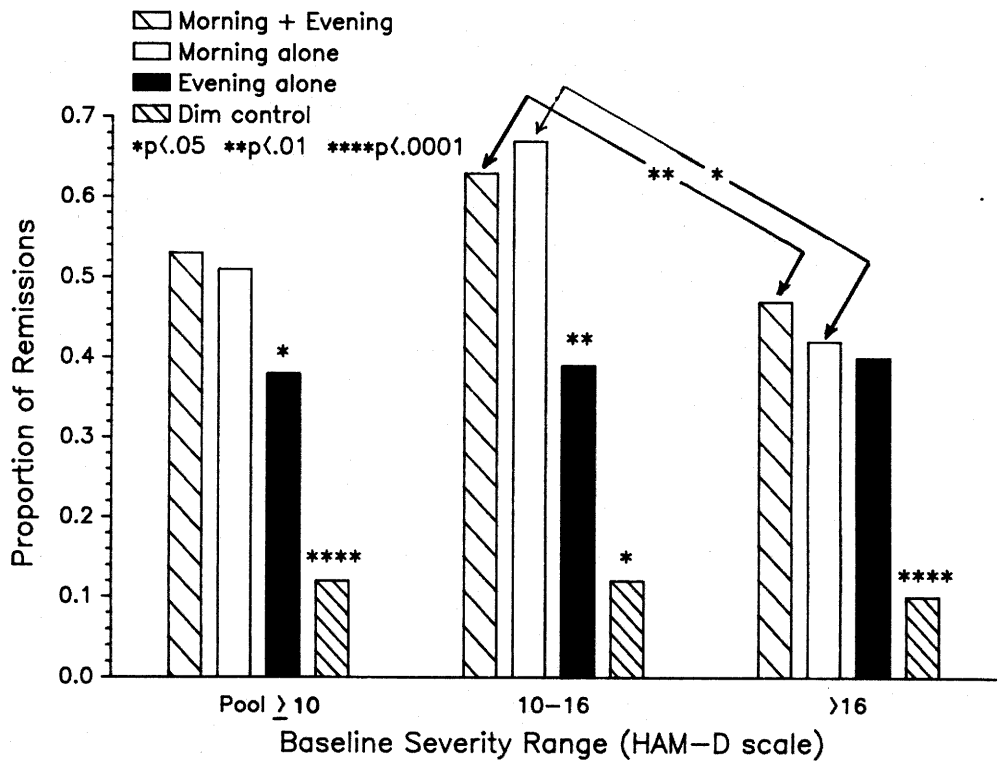



**Figure 3.** Proportion of patients achieving clinical remissions under light therapy, defined as a pre- to posttreatment HAM-D score reduction of at least 50% to a level below 8, as a function of baseline symptom severity. Pool of patients with baseline scores of 10 or higher is shown at left. Middle and right-hand sets of histograms divide the pool into mildly depressed (pretreatment HAM-D score of 10–16) and moderately-to-severely depressed (>16) subsets. Significant variations across groups were determined by hypothesis tests of differences in proportions.

vidual crossover studies have found no reliable morning–evening bright-light difference but did find quite respectable rates of remission for evening treatment (e.g., Terman et al. 1987; A. Wirz-Justice, personal communication). Still, we have argued that since exclusive morning responders far outnumber exclusive evening responders, nondifferential responders might be sensitive to some other aspect of the therapy. Lewy and Sack (1986) postulated a nonspecific energizing effect apart from specific therapeutic circadian phase advances obtained only in early morning. Light therapy studies have unfortunately not used a placebo washout prior to the acute treatment phase, and high initial expectations for improvement are likely to color the treatment response for the first several weeks. We have observed evening light to be markedly less effective than morning light when tested in subjects with substantial prior experience, whereas new subjects appear much more responsive. Testing the persistence of the evening treatment response in comparison to that of morning response would be an important step toward clarifying this issue. Benefits of a nonspecific treatment —in this situation, one that shows no dose dependency, no differential time-of-day effect, and low efficacy across the patient population—might wane. Studies varying the circadian phase of evening light, early or late, are also needed to establish whether later exposures are consistently countertherapeutic by virtue of phase delays.

Consideration of the midday results is limited by the number of subjects studied ($n$ = 34); it is notable, however, that two of three studies that used this regimen did achieve a level of efficacy approaching that of morning light. This may indicate that the midday responders were dose dependent and insensitive to time-of-day effects. Alternately, it may support the hypothesis of a delayed phase-response curve in SAD patients (Lewy et al. 1987), wherein a midday light exposure given early enough to cause a phase advance could induce a clinical change similar to that of morning light. If so, morning and midday procedures are thus not functionally distinct. However, it is also true that midday light yields a global proportion of remissions of 0.32, by the joint criteria, which might be considered to fall within a placebo response range, without significant difference from evening light in effectiveness.

For documented cases of phase advance at baseline, with early melatonin onset or early sleep phase (Lewy et al. 1987), however rare, evening light exposure would be expected to be therapeutically active

in long-term treatment and morning light ineffective. The clinician must face the difficulties of attempting a diagnosis of circadian phase—e.g., through analyses of melatonin, temperature, activity, and sleep patterns—before recommending a time of day for light treatment, or simply try both times to determine which is best in individual cases. Lewy (1987) and Terman (1988b) outlined various clinical strategies for administration of light therapy as related to putative circadian system interactions.

### Control Procedures

***Estimation of a Placebo Response Rate.*** Development of convincing placebo controls for this research presents a formidable problem. The active treatment cannot be administered blind, and dim or brief light exposure controls are not necessarily convincing to the patient as treatment alternatives. Nonetheless, even without direct estimation of a nonspecific effect, demonstration of differential response among light treatments can provide convincing evidence of efficacy. A placebo response rate has not been directly established for the SAD population. Ideally, we would like to see a light-treatment effect exceed 20% to 40% placebo response rates commonly obtained in medication studies. Our group has, however, studied nonseasonal depressives under placebo-controlled trials of tricyclic, tetracyclic, and monoamine oxidase inhibiting antidepressant medications. Many of these patients have shown atypical vegetative symptoms similar to those found in SAD. Across several studies (cf. Rabkin et al. 1987) we have found 10-day placebo response rates of approximately 20%, and 6-week rates of approximately 30%. Interestingly, 10-day placebo responders tended to have milder symptoms than 6-week responders and to show decreased response in the winter months (cf. Terman 1988a). Although the 10-day estimate provides a close match to the 1-week treatment duration of most light therapy studies, antidepressant medication response ordinarily takes much longer. The optimum treatment duration for lights has yet to be established.

Previous studies suggest that placebo responses in depressed patients occur more frequently among milder cases (Rabkin et al. 1987). Fairchild et al. (1986) found such a trend in Beck and Carroll self-rating scales, although not in the HAM-D scale. If the general finding also applied to SAD patients, one would expect to find higher response rates for all lighting conditions within the "mild" subset of Figure 3 than within the "moderate-to-severe" subset. Instead, there is a striking equivalence of response rates for evening light and for dim light (two ostensible placebo conditions) across severity ranges, with stronger clinical response restricted to mild cases treated with morning light. There may have been an interaction between season and baseline severity underlying heightened placebo response in earlier studies, given significantly higher summertime placebo rates (Rabkin et al. 1987) that, perforce, are absent among winter depressives. The selective advantage of morning light among mild SAD cases, rather than reflecting a placebo response, is more plausibly interpreted as a dose dependency: the standard intensity of morning light (2500 lux) may constitute a low treatment dose that offers specific clinical benefit mainly in milder cases (see Dosing Dimensions of Light, below).

A double-blind medication study of winter depression, using a conventional placebo control, showed a 4-week placebo response rate of 17%, by our joint criteria, in comparison to 78% response to *d*-fenfluramine administered to 18 carbohydrate cravers with extreme weight gain (D. O'Rourke, personal communication). Viewed within the context of our survey, this result suggests that direct serotoninergic stimulation can equal or exceed light therapy in clinical efficacy. The placebo response rate of 17% does not differ statistically from the 11% rate obtained for the dim light control but is lower than the 38% rate under evening light ($p < 0.05$), a result that argues for partial efficacy of evening light.

Under the phase-shift hypothesis of Lewy et al. (1988), a given patient should respond selectively to morning or evening light—to morning if phase delayed at baseline, to evening if phase advanced. By implication, patients showing nondifferential remissions to morning and evening light might be considered placebo responders. Morning–evening crossover studies (Fig. 1a) support this scenario, with 62% of patients showing remissions with morning light, and 28% showing remissions with evening light. The overlap of these two groups—those responding nondifferentially with both morning and evening light—comprise 25% of patients. Assuming that these patients are placebo responders, we would estimate that there were 62% − 25% = 37% true morning responders and 28% − 25% = 3% true evening responders in the group. The few true evening responders would be expected to show abnormally early nocturnal melatonin onsets and body temperature declines, accompanied by early bedtimes and rise times. Indeed, the one exclusive evening-light responder from Oregon (Fig. 1a) did show a phase-advanced melatonin onset at baseline (Lewy et al. 1988), and several other such cases have recently been identified (Avery et al. 1988; A. Lewy, personal communication). Until the mechanism of therapeutic action is determined, however, regarding nondifferential morning–evening response as a placebo effect must be considered tentative.

***Subtherapeutic Light Exposure.*** Two major manipulations of light exposure parameters, intensity and duration, have been studied as ostensible placebo controls (Table 2) on the premise that the treatment effect requires bright light (>2000 lux) presented in long sessions (≥2 hours). Six studies have crossed subjects over between standard bright-light presentations (2 to 6 hours total exposure per day) and a variety of dim light values (≤400 lux), which is within the normal range of interior room illumination. Although two of these studies obtained statistically significant improvement in HAM-D scores after dim light, the proportion of subjects meeting our joint remission criteria was low (Wirz-Justice et al. 1986, 0.16; R. Depue, personal communication, 0.0). The proportion of dim-light remissions for the pooled group, 0.11, falls significantly below the proportion under evening bright light, 0.38 ($p < 0.0001$). Thus, evening light appears effective in comparison. One problem with this conclusion, however, is that patients' expectations of treatment efficacy have tended to be lower for dim light (Rosenthal et al. 1984), thus lessening the likelihood of a placebo response.

Two of the larger studies to date ($n = 43$) have presented relatively brief bright-light exposures (30 minutes, morning alone or morning plus evening), surmising that positive treatment effects require longer durations. In both cases, reductions in HAM-D scores were highly significant, and yet the proportion of subjects meeting our joint criteria for remission still fell within the 30% range. Several patients did show convincing responses to brief light exposure, however, and settled upon the 30-minute regimen for long-term maintenance therapy. This suggests that duration is a continuous dimension under which some supersensitive patients will benefit from brief exposures. For the SAD population as a whole, however, at levels of illumination in the 2500-lux range, morning exposures of up to several hours will be required.

## Sample Sizes for Future Research

The variability of outcomes across light therapy studies is substantial and is probably attributable to small sample sizes, beyond any procedural differences and variations in clinical profiles and assessment methods. With 62% of subjects achieving essentially complete clinical remissions under morning light in comparison to 28% under evening light (see morning versus evening crossover analysis above), the expected effect size ($h = 0.70$, computed from differences in proportions meeting the joint remission criteria) would require samples of 25 in experimental and control groups in order to achieve a power of 0.80 at the $p < 0.05$ level. This meets the level of power usually recommended (Cohen 1977). The power analysis strongly argues that sample sizes of light therapy studies to date have been too small to permit demonstration of statistical differences between morning and evening light, if in fact they do exist.

## Treatment Assessment

Subjects' baseline symptom severity has varied widely both within and between light therapy studies. Indeed, mean pretreatment Hamilton depression scores have ranged between 13.0 and 30.0. Although subjects have all met standard criteria for seasonal recurrence of depressive symptoms (Rosenthal et al. 1984), this wide baseline range makes it unlikely that any single treatment regimen would suffice for the entire population. One source of baseline variability may lie in rating criteria for the Hamilton interview, as applied by different groups. Though the HAM-D scale has apparently worked well across a wide range of studies, comparability of results would be improved by use of a structured interview guide (Williams 1988). A serious drawback of the Hamilton instrument for workers in SAD is the absence of items assessing vegetative symptoms such as hypersomnia and increased appetite that are prominent among SAD patients, sometimes even outweighing the melancholic symptoms. Though the severity of vegetative symptoms has not been consistently reported by light therapy investigators, a scale rating these symptoms, devised by N.E. Rosenthal, has been incorporated into a structured clinical interview for SAD (SIGH-SAD; Williams et al. 1988). Our own work indicates that differential responses to various light treatments are accentuated when the atypical vegetative symptoms are considered (Terman et al. 1989b).

Although the first large-scale clinical trial (Rosenthal et al. 1984) utilized a treatment period of 2 weeks, 1-week assessments of acute treatment have been more common in these studies. These are brief in comparison to medication studies, but the response to light typically occurs within 3 or 4 days. The result, however, is that the posttreatment Hamilton interview often assesses only a few days of improvement, after which the patient is immediately withdrawn or crossed over to a new condition. This restricted time frame for the posttreatment assessment necessarily lowers the confidence with which symptoms can be scaled in comparison to standard full week assessments and probably contributes to the variability of reported scores. Thus, we suggest extending treatment trials to at least 10 to 14 days to allow for reliable assessments of treatment responses, including those of slower responders. Poin-

tedly, Depue's (personal communication) study of morning plus evening light, with evaluation after 2 weeks of light, achieved the highest effect size of any in the set ($d = 5.02$).

### Dosing Dimensions of Light

This review has concentrated on time-of-day effects for light administration. There are several lines of evidence, however, that other dimensions of light exposure may determine treatment efficacy. The duration of the treatment session, for example, as noted above, may be a dosing variable requiring individual adjustment for the patient. Checkley et al. (1986) found that 1-hour sessions in morning and evening were ineffective in a severely affected group with mean baseline HAM-D of 29.6, whereas 3-hour sessions yielded remissions in 45% of cases. It is reasonable to surmise, then, that longer exposures would be required to accurately assess response in more severe cases.

Increasing the duration of daily light exposure is perhaps the least attractive alternative, however, because it ties down the patient. The finding that 1 week of light therapy yields approximately equivalent relative improvement in HAM-D scores across the range of baseline severity, but brings few severely depressed patients into the asymptomatic range [i.e., HAM-D scores under 8 (Fig. 3)], suggests that another alternative would be to extend the length of treatment. Light exposure for more than 7 to 10 days might increase success rates among patients whose scores require greater proportional reduction to reach the asymptomatic range.

A third dosing strategy would be to increase light intensity. Our current work indicates that patients who respond to 2 hours of morning light at 2500 lux —but relapse at briefer durations—show an immediate positive response when given 10,000 lux for 30 minutes (Terman 1988a,b). The assessment of efficacy necessarily rests on an adequate dosing regimen, and 2500-lux studies may have been at a handicap.

## CONCLUSION

A survey of 29 light therapy studies reported between 1984 and 1987 shows considerable variability in treatment outcome as a function of time of day and duration of light administration. Data pooled across studies indicate a clear positive treatment effect for early morning bright light exposure and equivocal results for midday and evening exposures. The highest remission rates have been obtained for mildly depressed patients, a group that also shows a differential response favoring morning over evening light. Interpretation of the evening light effect—as truly antidepressant, transiently energizing, or within the range of placebo response—remains problematic. Evaluation of clinical efficacy may be improved by use of brighter lights, longer courses of treatment, systematic use of washouts within crossover designs, stratification of results on the dimension of baseline severity, a validated scale for assessment of atypical vegetative symptoms, and substantially larger sample sizes than have been used in research to date.

## ACKNOWLEDGMENTS

This research was supported by NIMH Grants KO2 MH00461, RO1 MH42931, and MHCRC 30906. We thank Donald Klein, Martha Link, Edward Nunes, Judith Rabkin, and David Schlager for comments on the manuscript; Jacob Cohen, Jeffrey Markowitz, Deborah Padgett, and Donald Ross for statistical consultations; and the participating investigators in each center for sharing raw data.

## REFERENCES

American Psychiatric Association (1987): Diagnostic and Statistical Manual of Mental Disorders, ed 3, revised (DSM-III-R). Washington, DC, American Psychiatric Association

Avery DH, Khan A, Dager SR, Dunner DL (1988): Winter depression and response to A.M. and P.M. light. Paper presented at American Psychiatric Association meetings, Montreal

Bassi CJ, Powers MK (1986): Daily fluctuations in the detectability of dim lights by humans. Physiol Behav 38:871–877

Bock RD (1975): Multivariate Statistical Methods in Behavioral Research. New York, McGraw-Hill

Bojkowski CJ, Arendt J (1988): Annual changes in 6-sulphatoxymelatonin excretion in man. Acta Endocrinol, in press

Bojkowski CJ, Aldhous ME, English J, Franey C, Poulton AL, Skene DJ, Arendt J (1987): Suppression of nocturnal plasma melatonin and 6-sulphatoxymelatonin by bright dim light in man. Horm Metab Res 19:437–440

Bowker AH (1948): A test for symmetry in contingency tables. J Am Stat Assoc 43:572–574

Brainard GC, Lewy AJ, Menaker M, Fredrickson RH, Miller LS, Weleber RG, Cassone V, Hudson D (1989): Dose-response relationship between light irradiance and the suppression of plasma melatonin in human volunteers. Brain Res, in press

Checkley S, Winton F, Franey C, Arendt J (1986): Effects of phototherapy upon mood and melatonin in seasonal affective disorder. Paper presented at Royal College of Psychiatry, London

Cochran WG (1963): Sampling Techniques, rev ed. New York, John Wiley and Sons

Cohen J (1977): Statistical Power Analysis for the Behavioral Sciences, rev ed. New York, Academic Press

DeCoursey PJ (1960): Phase control in a rodent. Cold Spring Harb Symp Quant Biol 25:49–55

Fairchild CJ, Rush AJ, Vasacada N, Giles DE, Khatami M (1986): Which depressions respond to placebo? Psychiatry Res 18:217–226

Ferguson GA (1959): Statistical Analysis in Psychology Education. New York, McGraw-Hill

Hamilton M (1967): Development of a rating scale for primary depressive illness. Br J Soc Clin Psychol 6:276–296

Hamilton M (1982): The effect of treatment on the melancholias. Br J Psychiatry 140:223–230

Hellekson CJ (1989): Phenomenology of Seasonal Affective Disorder. In Rosenthal NE, Blehar M (eds), Seasonal Affective Disorders and Phototherapy. New York, Guilford Press, in press

Hellekson CJ, Rosenthal NE (1986): New light on Alaskan "cabin fever." Paper presented at American Psychiatric Association meetings, Washington, DC

Hellekson CJ, Kline JA, Rosenthal NE (1986): Phototherapy for seasonal affective disorder in Alaska. Am J Psychiatry 143:1035–1037

Honma K, Honma S, Wada T (1987): Phase-dependent shift of free-running human circadian rhythms in response to a single bright light pulse. Experientia 43:1205–1207

Isaacs C, Stainer DS, Sensky TE, Moor S, Thompson C (1988): Phototherapy and its mechanisms of action in seasonal affective disorder. J Affect Disorders 14:13–19

Jacobsen FM, Rosenthal NE (1986): Seasonal affective disorder and the use of light as an antidepressant. Directions in Psychiatry 6:1–8 (Lesson 3)

Jacobsen FM, Wehr TA, Skwerer RA, Sack D, Rosenthal NE (1987): Morning versus midday phototherapy of seasonal affective disorder. Am J Psychiatry 144:1301–1305

James SP, Wehr TA, Sack DA, Parry BL, Rosenthal NE (1985): Treatment of seasonal affective disorder with evening light. Br J Psychiatry 147:424–428

Kasper S, Rogers S, Yancey A, Skwerer RG, Schulz PM, Rosenthal NE (1989): Psychological effects of light therapy in normals. In Rosenthal NE, Blehar M (eds), Seasonal Affective Disorders and Phototherapy. New York, Guilford Press, in press

Klein DF, Gittleman R, Quitkin FM, Rifkin A (1980): Diagnosis and Drug Treatment of Psychiatric Disorders. Baltimore, Williams & Wilkins

Lewy AJ (1987): Treating chronobiological sleep and mood disorders with bright light. Psychiatr Ann 17:664–669

Lewy AJ, Sack RL (1986): Light therapy and psychiatry. Proc Soc Exp Biol Med 183:11–18

Lewy AJ, Wehr TA, Goodwin FK, Newsome DA, Markey SP (1980): Light suppresses melatonin secretion in humans. Science 210:1267–1269

Lewy AJ, Wehr TA, Goodwin FK, Newsome DA, Rosenthal NE (1981): Manic-depressive patients may be supersensitive to light. Lancet 1:383–384

Lewy AJ, Kern HE, Rosenthal NE, Wehr TA (1982): Bright artificial light treatment of a manic-depressive patient with a seasonal mood cycle. Am J Psychiatry 139:1496–1498

Lewy AJ, Sack R, Singer CM (1985): Treating phase typed chronobiologic sleep and mood disorders using appropriately timed bright artificial light. Psychopharmacol Bull 21:368–372

Lewy AJ, Sack RL, Miller S, Hoban TM (1987): Antidepressant and circadian phase-shifting effects of light. Science 235:352–354

Lewy AJ, Sack RL, Singer CM, White DM (1988): The phase shift hypothesis for bright light's therapeutic mechanism of action: Theoretical considerations and experimental evidence. Psychopharmacol Bull 23:349–353

Light RJ, Smith PV (1971): Accumulating evidence: Procedures for resolving contradictions among different research studies. Harvard Educ Rev 41:429–471

Lingjaerde O, Bratlid T, Hansen T, Gøtestam KG (1986): Seasonal affective disorder and mid-winter insomnia in the far north: Studies on two related chronobiological disorders in Norway. Proc Coll Int Neuro-Psychopharmacol 15:187–189

Murphy DGM, Abas M, Franey C, Arendt J, Binnie C, Checkley S (1989): Seasonal affective disorder: A neurophysiological approach. In Thompson C, Silverstone T (eds), Seasonal Affective Disorder. London, CRC Clinical Neuroscience, in press

Potkin S, Zetin M, Stamenkovic V, Kripke DF, Bunney WE Jr (1986): Seasonal affective disorder: Prevalence varies with latitude and climate. Clin Neuropharmacol 9(Suppl 4):181–183

Quitkin FM, Rabkin JG (1981): Methodological problems in studies of depressive disorder: Utility of the discontinuation design. J Clin Psychopharmacol 1:283–288

Rabkin JG, Stewart JW, McGrath PG, Markowitz JS, Harrison W, Quitkin FM (1987): Baseline characteristics of ten-day placebo washout responders in antidepressant trials. Psychiatry Res 21:9–22

Rosenthal NE, Sack DA, Gillin JC, Lewy AJ, Goodwin FK, Davenport Y, Mueller PS, Newsome DA, Wehr TA (1984): Seasonal affective disorder: A description of the syndrome and preliminary findings with light therapy. Arch Gen Psychiatry 41:72–80

Rosenthal NE, Sack DA, Carpenter CJ, Parry BL, Mendelson WB, Wehr TA (1985a): Antidepressant effects of light in seasonal affective disorder. Am J Psychiatry 142:606–608

Rosenthal NE, Sack DA, James SP, Parry BL, Mendelson WB, Tamarkin L, Wehr TA (1985b): Seasonal affective disorder and phototherapy. Ann New York Acad Sci 453:260–269

Rosenthal NE, Skwerer RG, Sack DA, Duncan CC, Jacobsen FM, Tamarkin L, Wehr TA (1987): Biological effects of morning-plus-evening bright light treatment of seasonal affective disorder. Psychopharmacol Bull 23:364–369

Rosenthal NE, Jacobsen FM, Sack DA, Arendt J, James SP, Parry BL, Wehr TA (1988a): Atenolol in seasonal affective disorder: A test of the melatonin hypothesis. Am J Psychiatry 145:52–56

Rosenthal NE, Sack DA, Skwerer RG, Jacobsen FM, Wehr TA (1988b): Phototherapy of seasonal affective disorder. J Biol Rhythms 3:101–120

Rosenthal NE, Targum SD, Docherty TP, Hoffmann HA, Hamovit JR, Bryant MJ, Kasper SF (1988c): Prevalence of SAD and S-SAD by latitude in continental United States. Paper presented at American Psychiatric Association meetings, Montreal

Sack RL, Lewy AJ, White DM, Singer CM, Hoban TM (1987): Morning light treatment for winter depression. Paper presented at American Psychiatric Association meetings, Chicago

Siegel S (1956): Nonparametric Statistics for the Behavioral Sciences. New York, McGraw-Hill

Skwerer RG, Duncan CC, Sack DA, Jacobsen FM, Tamarkin L, Wehr TA, Rosenthal NE (1988): Neurobiology of seasonal affective disorder and phototherapy. J Biol Rhythms 3:135–154

Spitzer RL, Endicott J, Robbins E (1978): Research diagnostic criteria: rationale and reliability. Arch Gen Psychiatry 35:773–782

Strassman RJ, Peake GT, Qualls CR, Lisansky EJ (1987): A model for the study of the acute effects of melatonin in man. J Clin Endocrinol Metabol 65:847–852

Terman M (1988a): On the question of mechanism in phototherapy for seasonal affective disorder: Considerations of clinical efficacy and epidemiology. J Biol Rhythms 3:155–172

Terman M (1988b): Light therapy. In Kryger M (ed), Principles and Practice of Sleep Medicine. Philadelphia, Saunders, pp 717–722

Terman M, Terman JS (1985): A circadian pacemaker for visual sensitivity? Ann NY Acad Sci 453:147–161

Terman M, Quitkin FM, Terman JS, Stewart JW, McGrath PJ (1987): Light therapy: Exposure duration and time of day. Paper presented at American Psychiatric Association meetings, Chicago

Terman M, Terman JS, Quitkin FM, Cooper TB, Lo ES, Gorman JM, Stewart JW, McGrath PJ (1988a): Response of the melatonin cycle to phototherapy for seasonal affective disorder. J Neur Transm 72:147–165

Terman M, Terman JS, Schlager D, Quitkin FM (1988b): Efficacy of 10,000 lux light therapy. Paper presented at World Psychiatric Association meetings, Washington, DC

Terman M, Botticelli SR, Link BG, Link MJ, Quitkin FM, Hardin TE, Rosenthal NE (1989a): Seasonal symptom patterns in New York: Patients and population. In Thompson C, Silverstone T (eds), Seasonal Affective Disorder. CRC Clinical Neuroscience, London, in press

Terman M, Terman JS, Quitkin FM, Stewart JW, McGrath PJ, Nunes EV, Wager SG, Tricamo E (1989b): Dosing dimensions of light therapy: Duration and time of day. In Thompson C, Silverstone T (eds), Seasonal Affective Disorder. CRC Clinical Neuroscience, London, in press

Thompson C, Isaacs C (1988): Seasonal affective disorder —a British sample: Symptomatology in relation to mode of referral and diagnostic subtype. J Affect Disorders 14:1–11

Wehr TA, Jacobsen FM, Sack DA, Arendt J, Tamarkin L, Rosenthal NE (1986a): Phototherapy of seasonal affective disorder: Time of day and suppression of melatonin are not critical for antidepressant effects. Arch Gen Psychiatry 43:870–875

Wehr TA, Skwerer R, Jacobsen FM, Sack DA, Rosenthal NE (1986b): Eye versus skin phototherapy of seasonal affective disorder. Am J Psychiatry 144:753–757

Wilcox RR (1987): New Statistical Procedures for the Social Sciences. Hillsdale, NJ, Lawrence Erlbaum

Williams JBW (1988): The SIGH-D: An interview guide for the Hamilton Depression Rating Scale. Arch Gen Psychiatry 45:742–747

Williams JBW, Link MJ, Rosenthal NE, Terman M (1988): Structured Interview Guide for the Hamilton Depression Scale—Seasonal Affective Disorder version (SIGH-SAD). New York State Psychiatric Institute

Wirz-Justice A, Bucheli C, Graw P, Kielholz P, Fisch H-U, Woggon B (1986): Light treatment of seasonal affective disorder in Switzerland. Acta Psychiatr Scand 74:193–204

Wirz-Justice A, Schmid AC, Graw P, Kraeuchi K, Poeldinger W, Fisch H-U, Buddeberg C (1987): Dose relationships of morning bright white light in seasonal affective disorders. Experientia 43:574–576

Wurtman RJ, Wurtman JJ (1989): Carbohydrates and depression. Sci Amer (Jan.), pp. 68–75

Yerevanian BI, Anderson JL, Grota LJ, Bray M (1986): Effects of bright incandescent light on seasonal and nonseasonal major depressive disorder. Psychiatry Res 18:355–364

Yerevanian BI (1987): Treatment of seasonal affective disorder with bright incandescent light: Is timing of treatment important? Paper presented at American Psychiatric Association meetings, Chicago