ON THE EXPLANATION OF SHORT REM LATENCIES IN DEPRESSION

VANDENHOOFDAKKER, RH; BEERSMA, DGM

Published in:
Psychiatry Research

DOI:
10.1016/0165-1781(85)90009-5

Psychiatry Research, **16**, 155-163
Elsevier

On the Explanation of Short REM Latencies in Depression

Rutger H. van den Hoofdakker and Domien G.M. Beersma

Abstract. Sleep in depression is characterized by the occurrence of episodes of rapid eye movement (REM) sleep at sleep onset. The empirical foundations of three hypotheses about the origin of this phenomenon are examined: (1) A circadian rhythm hypothesis stating that sleep onset REM episodes (SOREMs) are the result of an abnormal phase-position of the REM sleep production cycle. (2) A REM sleep-slow wave sleep interaction hypothesis that attributes SOREMs to a low non-REM sleep propensity. (3) A circadian amplitude hypothesis, in which a flattening of the circadian arousal cycle is thought to be causally related to SOREMs. None of the hypotheses are found to be supported by firm empirical evidence.

Key Words. Endogenous depression, REM sleep latency, circadian rhythms, ultradian rhythms, sleep models.

Short latencies of rapid eye movement (REM) sleep are observed in a large variety of conditions (see Schulz and Lund, 1985). In some of these conditions it has been shown that the distribution of REM latencies is bimodal; i.e., the first REM sleep episode begins either within 20 minutes (the sleep onset REM episode, or SOREM) or after about 60-70 minutes. Schulz and Lund (1985) have suggested that this bimodal distribution is typical for all cases in which a very short mean REM latency has been observed. Whether this generalization is justified or not is an important issue with implications for the understanding of the regulation of the non-REM/REM cycle and for pathophysiological theories of depression. However, Schulz and Lund (1985) confined their examination largely to the field of depression where the bimodal distribution of REM latencies is well documented. In this article, we add to Schulz and Lund's elegant discussion of the hypotheses and the empirical evidence involved, making use of data obtained in a study from our own laboratory.

Methods

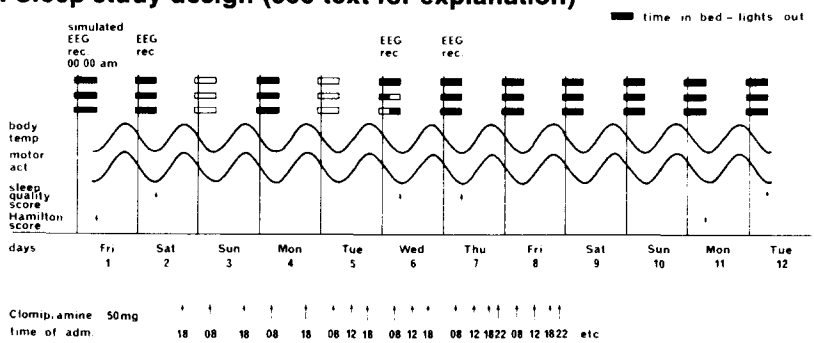
Relevant data were obtained in a study designed for other purposes (Elsenga and van den Hoofdakker, 1982). The design is shown in Fig. 1. Three groups of 10 endogenously depressed patients, diagnosed by two independent psychiatrists, were studied.

Earlier versions of this article and the following article by Brambilla et al. were presented as part of a symposium at the World Congress of Psychiatry, Vienna, July 1983. Both articles were submitted on January 18, 1984, and revised October 1, 1984, and November 9, 1984, respectively. They are part of a group of articles on the "Psychobiology of Depression," assembled by Prof. Dr. D. von Zerssen and focusing on chronobiological and psychoendocrine research on depressive disorders. Other articles from this symposium appeared in the August and September issues.

Rutger H. van den Hoofdakker, M.D., and Domien G.M. Beersma, Ph.D., are in the Department of Biological Psychiatry, University Psychiatric Clinic, Oostersingel 59, 9713 EZ Groningen, The Netherlands. (Reprint requests to Prof. Dr. R.H. van den Hoofdakker.)

0165-1781/85 $03.30

Fig. 1. Sleep study design (see text for explanation)



Group I. Group I comprised seven women and three men with a mean age of 51 years. Their mean score on the Hamilton Rating Scale for Depression (HRSD) was 32. Group I had the following sleep schedule: night 1, ad libitum sleep (ALS) between 2300h and 0800h; night 2, total sleep deprivation (TD); night 3 ALS; night 4 TD; nights 5-11 ALS between 2300h and 0800h.

Group II. Group II comprised seven women and three men with a mean age of 50 years. Their mean score on the HRSD was 31. Group II followed the same schedule as Group I except on night 5. Then, 3 hours of sleep were allowed between 2300h and 0300h.

Group III. Group III comprised six women and four men with a mean age of 48 years. Their mean score on the HRSD was 35. Group III also followed the same schedule but on night 5, they were allowed 3 hours of sleep between 0400h and 0800h.

Measurement. After one adaptation sham recording (night 0), sleep was recorded on nights 1, 5, and 6. During the total experimental period, rectal temperature was recorded continuously. A large number of other measurements were made daily (Elsenga and van den Hoofdakker, 1982; Beersma et al., 1983*a*, 1983*b*), but these data are not relevant to the present report.

The patients were drug free for at least 3.5 days before the experiment began. During the experimental period, they received clomipramine (Cl) according to the schedule presented in Fig. 1. Plasma levels of Cl and desmethyl Cl were assessed on days 4, 5, 8, and 10. No significant differences were found between the groups.

Although 37 patients started the study, only 30 completed it. Differences in numbers in this report reflect missing data due to dropouts or problems of a technical nature. Results are presented as mean ± SD.

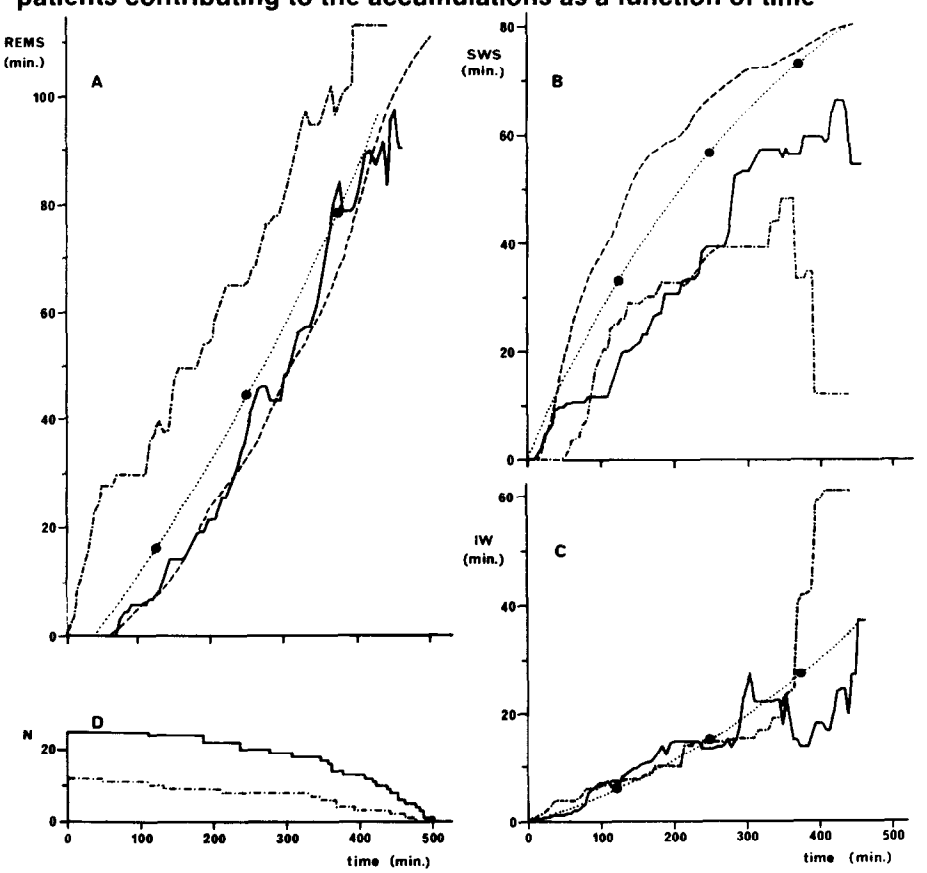
Circadian Rhythm Hypothesis.

The first hypothesis discussed by Schulz and Lund (1985) states that the SOREMs in the sleep of depressed patients are the result or the expression of a phase-advance of the circadian REM sleep rhythm with respect to the sleep-wake rhythm. Since in healthy subjects the REM sleep rhythm covaries with the rhythms of body temperature and cortisol secretion, the existence of a phase-advance of these latter rhythms in depressed patients might strengthen the hypothesis.

The critical question is: What is the empirical evidence for such an abnormal phase-position of these rhythms in depression? The data collected in the experiment described above provide information on the phase-positions of REM sleep production and temperature.

REM Sleep. If the circadian rhythm of REM sleep production in depression were phase-shifted with respect to the sleep-wake rhythm, one would expect a difference in the course of REM sleep production during sleep between depressed patients and healthy controls. We examined this question in the following way. In the 37 baseline nights of our experiment (night 1 in Fig. 1), a clear-cut bimodal distribution of REM latencies was found, with 12 patients showing latencies of 5.5 ± 5 minutes, and the remaining 25 of 88 ± 52 minutes. Obviously, depressed patients begin sleep in two distinct ways. In order to examine how they produce REM sleep during the total sleep period, the rates at which REM sleep accumulated after sleep onset were calculated. Fig. 2a shows the median accumulation after sleep onset during SOREM and non-SOREM sleep. Sleep onset times were not significantly different: 01:04 ± 101 minutes and 00:36 ± 52 minutes in SOREM and non-SOREM sleep, respectively. The

Fig. 2. Accumulation of REM sleep (A), slow wave sleep (B), and intermittent wakefulness (C) during drug-free baseline sleep, and (D) numbers of patients contributing to the accumulations as a function of time

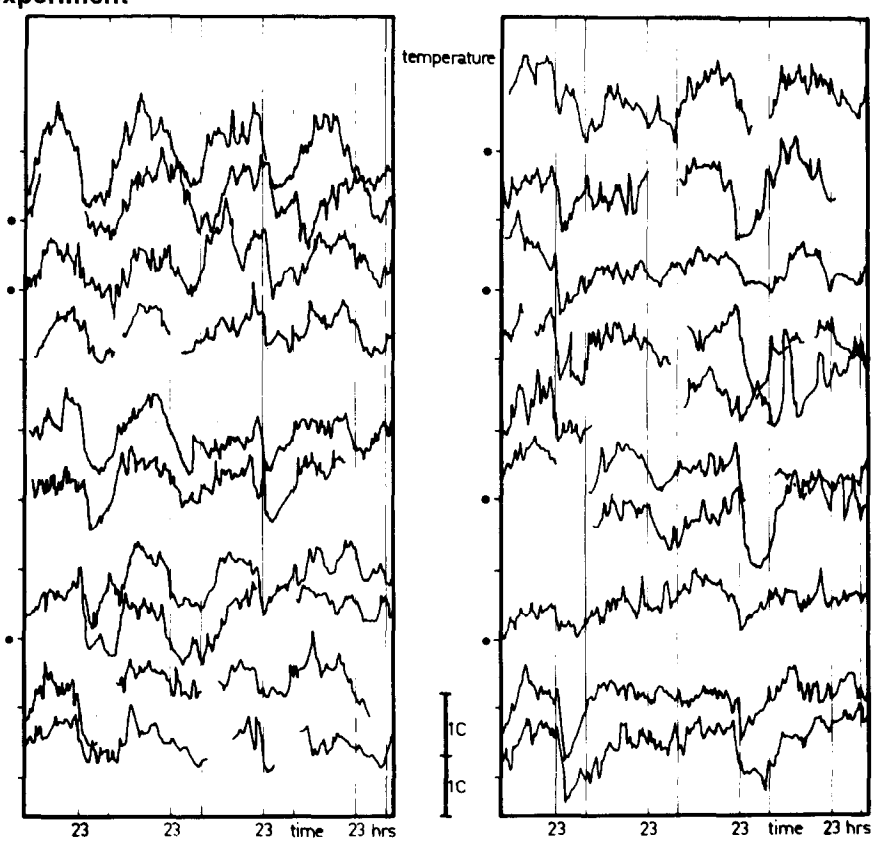


Solid curves = median accumulations in 25 nights with REM latencies of 88 ± SD 52 minutes; dashed and dotted curves = median accumulations in 12 nights with REM latencies of 5.5 ± SD 5 minutes; dashed curves = accumulations obtained by Gaillard (1977) and Gaillard and Martin (1975); dotted lines = accumulations calculated on the basis of data from Spiegel (1981).

difference in accumulation is restricted to the first 40 minutes of sleep; thereafter, REM sleep follows a similar course during the night, as demonstrated by the close parallelism of the curves. Fig. 2a contains two additional accumulation curves that allow REM sleep rates in depressed patients and controls to be compared. One curve is derived from Gaillard and Martin (1975) and shows the course in 13 healthy young men; the other was calculated from Spiegel's (1981) data on 57 healthy elderly subjects. Apparently, the course of REM sleep in our patients is quite normal, except for a brief disturbance after sleep onset in about 30% of the nights (see also Beersma et al., 1983*b*).

Body Temperature. Is there any evidence of a phase-advance of the body temperature rhythm in depressed patients? Fig. 3 shows those curves that are of sufficient quality to provide an impression of the course of body temperature between 0900h on day 1 and 0900h on day 5. Comparison of the ALS nights with the TD nights shows the strong masking influence of sleep. In many TD nights circadian modulation

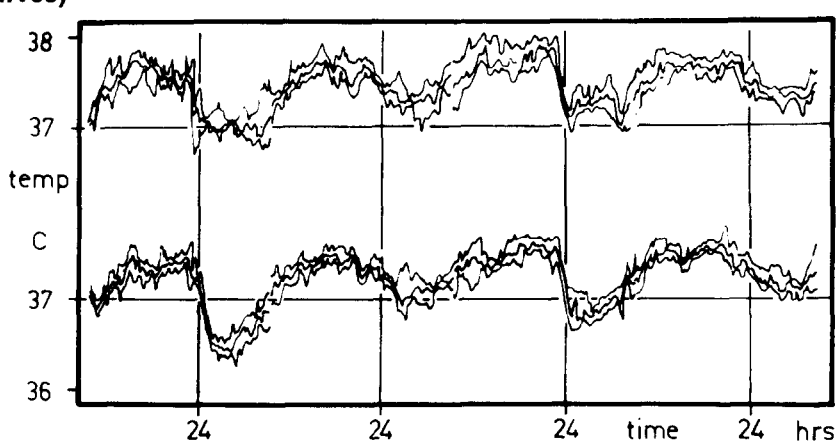
Fig. 3. Course of body temperature of 20 patients during the first 4 days of the experiment



For each curve, the 37C level is indicated on the abscissa. Curves of patients who showed a SOREM in night 1 are indicated by a dot. In nights 1 and 3 ad libitum sleep was permitted, whereas patients were sleep-deprived during nights 2 and 4.

is faint or even indetectable. Thus, under normal sleep-wake conditions, no reliable estimate can be obtained about the phase-position of the endogenous temperature rhythm. Both manifest abnormal and normal positions of temperature minima may be the result of masking. A better indication of the presence or absence of abnormalities is provided by the "unmasked" parts of the temperature curve, i.e., the parts recorded during TD. In those cases in which significant circadian modulation exists, no convincing indications of a phase-advance are found; the majority of the detectable minima are located in the later part of the night. In Fig. 4, the median temperature values as a function of time are presented for the patients with (upper curves) and without SOREMs (lower curves). (We discuss these curves in greater detail below.) No obvious phase-shift could be detected, either in the SOREM or in the non-SOREM group.

Fig. 4. Median body temperature values as a function of time for 7 patients who showed a SOREM in night 1 (upper curves) and 13 who did not (lower curves)



The median ± 1 curves are also plotted. The distances from these curves to the median curves represent a fair estimate of the standard errors of the mean. Sleep schedule is as in Fig. 3.

REM Sleep-Slow Wave Sleep (SWS) Interaction Hypothesis

Borbély's (1982) model of the regulation of sleep and Borbély and Wirz-Justice's (1982) model of depression provide the theoretical basis of this hypothesis. Timing, duration, and intensity of non-REM sleep are thought to be controlled by the interaction of a homeostatic process S and a circadian process C. The hypothetical process S increases as the duration of wakefulness increases, but it decreases during sleep. Thus, S represents a "need" for non-REM sleep, the time course of which is assumed to be reflected in the course of electroencephalographic (EEG) power density. Process C represents the circadian variation of the "sleep threshold." Production of non-REM sleep increases when S increases and when C decreases, and vice versa. The "sleep threshold" process C may influence non-REM sleep production in two ways: by setting circadian limits to total sleep duration and by sleep fragmentation as a consequence of transient fluctuations. The amount of REM sleep is

controlled by an ultradian mechanism in which need for non-REM sleep and REM sleep interact reciprocally. The formal properties of the mechanisms are based on McCarley and Hobson's (1975) model of interaction between REM sleep-enhancing and REM sleep-inhibiting neuronal populations. As to the content of the model, the interaction between these neuronal populations is replaced by the interaction between REM sleep and non-REM sleep need, respectively.

According to Borbély and Wirz-Justice (1982), process S is deficient in depression. This would explain the low non-REM sleep propensity, as reflected by the low production of stages 3 and 4 (SWS), which would, in turn, account for the short REM latency.

Schulz and Lund (1975) conclude that if short REM sleep latencies were caused by a lowered non-REM sleep propensity, an intrapatient correlation between amounts of non-REM sleep (especially SWS) and REM sleep latency would be expected. In the absence of such a correlation, they rejected the hypothesis. We tested the hypothesis in a slightly different way. If the deficit in process S were causally related to the occurrence of SOREMs, the production rate of SWS, supposedly a reflection of process S, would be expected to be smaller in SOREM sleep than in non-SOREM sleep. For comparison of accumulation curves of SWS in SOREM and non-SOREM sleep, see Fig. 2b. Evidently, the accumulation of SWS is roughly the same in the first 4 hours of SOREM and non-SOREM sleep (the apparent, but nonsignificant differences in median SWS accumulation during sleep at the end of the night are due to the small number of patients contributing to this part of the curve). Clearly, the data do not support the REM-SWS interaction hypothesis.

One further remark regards the two other curves in Fig. 2b. They demonstrate SWS accumulation in the healthy young and elderly subjects examined by Gaillard (1977) and Spiegel (1981). The differences between our depressed patients and the healthy subjects are consistent with many observations that SWS is deficient in depression.

The arguments against the REM-SWS hypothesis are based on the assumption that SWS is an adequate indicator of process S. However, this assumption is questionable. The time spent in SWS undoubtedly represents a measure of EEG power density, being a rough estimation of the integration of power during those parts of the sleep records in which it peaks. As stated above, in Borbély's model, non-REM propensity (and thus SWS propensity or power density) is not determined by process S, but by the interaction of process S and process C. Thus, a diminished production of SWS as demonstrated in Fig. 2b does not necessarily reflect an insufficient need for non-REM sleep, i.e., of a deficient process S. Changes in process C may be responsible as well. A more detailed discussion of this issue can be found in Beersma et al. (in press). The evidence for a deficiency of the need for non-REM sleep in depression is somewhat weaker than the low production of SWS suggests (van den Hoofdakker and Beersma, 1984). Consequently, the failure to find a relation between SWS and REM latency may be not surprising.

Circadian Amplitude Hypothesis

This hypothesis is an alternative to the former two. Schulz and Lund (1985) propose that flattening of the circadian arousal cycle or flattening of its measurable indicators

increases the probability for the occurrence of SOREM episodes. As the temperature rhythm is taken as such an indicator, the more specific assumption is: A high probability for SOREMs exists in those persons who have a small day-night temperature difference.

The main argument in favor of this hypothesis is the finding by Schulz and Lund (1985) that both depressed and healthy subjects who produce SOREMs show a significantly smaller amplitude of the educed body temperature curves than their counterparts who do not (see also Schulz and Lund, 1983).

A comparison of temperature amplitudes in patients who did and did not produce SOREMs can be obtained by inspecting the curves in Fig. 4. It should be noted that SOREMs occurred only in night 1 because REM sleep was completely or almost completely suppressed by clomipramine in night 3 and sleep deprivation occurred in nights 2 and 4. Two differences can be detected in the first 24 hours between the upper curve and the lower curve: (1) Patients with SOREMs have a smaller circadian amplitude, due to a smaller decrease during the night. (2) These patients show a higher circadian mean. The first observation seems to be in accord with the findings of Schulz and Lund and, consequently, to support the hypothesis that SOREMs are facilitated by a flattening of the circadian temperature rhythm. However, like Fig. 3, Fig. 4 reveals that sleep has an enormous masking influence; the circadian modulation in those nights in which ALS was allowed (nights 1 and 3) is far greater than in TD nights (nights 2 and 4). Therefore, the contribution of the endogenous factor to the shape of the curves during night 1 is very uncertain. The portions of the curves obtained during TD nights may be more representative of the endogenous variation. Comparison of these portions shows no differences between the SOREM and the non-SOREM groups. This observation strongly suggests that the difference between the body temperature curves in SOREM and non-SOREM nights is a consequence of the difference in sleep architecture rather than its cause.

The second observation, i.e., the difference in mean circadian temperature between the groups, also requires comment. As Schulz and Lund (1985) point out, the body temperature cycle may be regarded as an indicator of a (hypothetical) arousal cycle. It might be the level of the arousal cycle rather than its amplitude that facilitates SOREMs. This assumption finds some support in a recent report of Asnis et al. (1983), according to which REM latency is inversely related to the level of plasma cortisol, another potential indicator of arousal.

One final comment on the high mean circadian temperature in SOREM nights regards the model of sleep regulation proposed by Borbély (1982), Daan and Beersma (1983), and Daan et al. (1984). As Beersma et al. (in press) show, sleep length is shortened by raising the level of process C, especially because it induces early morning awakening. On the assumption that body temperature is related to process C, their theoretical proposal may be corroborated by our findings that the total sleep period in the patients who produced SOREMs is shorter. SOREM nights lasted 307 ± 147 (SD) minutes as compared to 369 ± 110 (SD) minutes for non-SOREM nights.[1] Which factors are responsible for this shortening? First, sleep has not been curtailed by a later

1. Sleep onset is defined by the first epoch of stage 2, 3, 4, or REM sleep that is followed by 10 minutes of sleep showing at most 2 minutes of intermittent wakefulness or stage 1 sleep. The end of sleep is defined by the last epoch of stage 2, 3, 4, or REM sleep preceded by 10 minutes of sleep showing at most 2 minutes of intermittent wakefulness or stage 1 sleep.

onset, because we showed that sleep onset times did not differ in patients with and without SOREMs, i.e., with relatively high and low temperature, respectively. A second possibility may be a large amount of intermittent wakefulness. Fig. 2c shows the accumulation curves of *intermittent* wakefulness in SOREM and non-SOREM sleep. There are no clear differences in the first 6 hours of sleep. Thus, the sleep of patients with higher body temperature has been curtailed by early morning awakening, again suggesting that the level of process C is of decisive importance to sleep architecture.

Conclusion

With respect to the circadian rhythm hypothesis, we agree with Schulz and Lund (1985) that SOREMs in depression cannot be explained as a result of a phase-advance of the REM sleep rhythm and other rhythms (e.g., body temperature) against the rest-activity cycle. Our agreement, however, is based on different arguments. Whereas Schulz and Lund (1985) seem to accept the existence of a phase-advanced REM sleep rhythm, we do not. In our opinion, even in sleep with a SOREM period, REM sleep is not phase-advanced, either against body temperature or against the sleep-wake cycle. In the absence of evidence for the existence of phase-disturbance, there is no reason to look for explanations of SOREMs in this direction.

We also agree with Schulz and Lund (1985) that little evidence exists in favor of the REM sleep-SWS interaction hypothesis. However, the REM sleep-SWS interaction hypothesis is a modification of Borbély and Wirz-Justice's (1982) original hypothesis, in which process S interacts with REM sleep production. Rejection of the first hypothesis does not imply rejection of the second. Testing the latter awaits specification of the parameters characterizing S at sleep onset.

We are skeptical about Schulz and Lund's (1985) alternative explanation of SOREMs: the circadian amplitude hypothesis. We did not find convincing evidence in support of the hypothesis that SOREMs are the consequence of a reduced amplitude of a circadian arousal rhythm. In contrast, our data suggest that the reduction of the amplitude of the temperature curves in SOREM nights is a consequence of a different influence of sleep on body temperature. There may be circadian factors that are important in the regulation of REM latencies, but they are more likely to be related to the level than to the amplitude of an arousal cycle.

A better explanation of SOREMs is provided by the mathematical model of McCarley and Hobson (1975). Beersma et al. (1984*a*, 1984*b*) simulated the unimodal distributions of REM sleep latencies in healthy elderly persons, as reported by Spiegel (1981), and the bimodal distribution of REM sleep latencies in depressives, as reported by Coble et al. (1981), on the basis of this model. The finding of a physiological basis of the model in humans will be the next step in elucidating the origin of SOREMs. Whether it will be found in sleep measures or in arousal measures is an open question.

References

Asnis, G.M., Halbreich, U., Sachar, E.J., Nathan, R.S., Ostrow, L.C., Novacenko, H., Davis, M., Endicott, J., and Puig-Antich, J. Plasma cortisol secretion and REM period latency in adult endogenous depression. *American Journal of Psychiatry*, **140**, 750 (1983).

Beersma, D.G.M., Daan, S., and van den Hoofdakker, R.H. Distribution of REM latencies and other sleep phenomena in depression as explained by a single ultradian rhythm disturbance. *Sleep,* **7,** 126 (1984*a*).

Beersma, D.G.M., Daan, S., and van den Hoofdakker, R.H. Sleep structure in depression. *Experimental Brain Research,* Suppl. **8,** 285 (1984*b*).

Beersma, D.G.M., Daan, S., and van den Hoofdakker, R.H. The timing of sleep in depression—Theoretical considerations *Psychiatry Research* (in press).

Beersma, D.G.M., van den Hoofdakker, R.H., Daan, S., and van Berkestijn, J.W.B.M. REM sleep and depression. In: Koella, W.P., ed. *Sleep 1982.* S. Karger, Basel, p. 349 (1983*a*).

Beersma, D.G.M., van den Hoofdakker, R.H., and van Berkestijn, J.W.B.M. Circadian rhythms in affective disorders: Body temperature and sleep physiology in endogenous depressives. In: van Praag, H.M., and Mendlewicz, J., eds. *Advances in Biological Psychiatry.* Vol. 11. S. Karger, Basel, p. 114 (1983*b*).

Borbély, A.A. A two process model of sleep regulation. *Human Neurobiology,* **1,** 195 (1982).

Borbély, A.A., and Wirz-Justice, A. Sleep, sleep deprivation and depression. *Human Neurobiology,* **1,** 205 (1982).

Coble, P.A., Kupfer, D.J., and Shaw, D.H. Distribution of REM latency in depression. *Biological Psychiatry,* **16,** 453 (1981).

Daan, S., and Beersma, D.G.M. Circadian gating of human sleep and wakefulness. In: Moore-Ede, M.C., ed. *Mathematical Modeling of Circadian Systems.* Raven Press, New York (1983).

Daan, S., Beersma, D.G.M., and Borbély, A.A. The timing of human sleep: Recovery process gated by a circadian pacemaker. *American Journal of Physiology,* **246,** R161 (1984).

Elsenga, S., and van den Hoofdakker, R.H. Total and partial sleep deprivation in endogenous depression. Presented at the 6th European Congress of Sleep Research, Zürich (1982).

Gaillard, J.M. Les tendances générales des stades du sommeil étudiées par ajustements de polynomes orthogonaux. *Electroencephalography and Clinical Neurophysiology,* **42,** 847 (1977).

Gaillard, J.M., and Martin, L. Fitting of a parabolic trend to the general component of cumulative occurrence of REM sleep in man. In: Koella, W.P., ed. *Sleep 1974.* S. Karger, Basel (1975).

McCarley, R.W., and Hobson, J.A. Neuronal excitability modulation over the sleep cycle: A structural and mathematical model. *Science,* **189,** 58 (1975).

Schulz, H., and Lund, R. Sleep onset REM episodes are associated with circadian parameters of body temperature: A study in depressed patients and normal controls. *Biological Psychiatry,* **18,** 1411 (1983).

Schulz, H., and Lund, R. On the origin of early REM sleep episodes in the sleep of depressed patients: A comparison of three hypotheses. *Psychiatry Research,* **16,** 65 (1985).

Spiegel, R. Sleep and sleeplessness in advanced age. In: Weitzman, E.D., ed. *Advances in Sleep Research.* Vol. 5. MTP Press Ltd., Lancaster (1981).

Van den Hoofdakker, R.H., and Beersma, D.G.M. Sleep deprivation, mood and sleep physiology. *Experimental Brain Research,* Suppl. **8,** 297 (1984).