Sleep, 3(1):83–92

Short Report

Wrist-Actigraphic Estimation of Sleep Time

D. J. Mullaney, D. F. Kripke, and S. Messin

University of California, San Diego, and Veterans Administration Medical Center, San Diego, California

Summary: Using a piezoelectric transducer, wrist activity was recorded simultaneously with electroencephalogram (EEG), electro-oculogram (EOG), and submental electromyogram (EMG) to obtain 102 recordings—39 from hospital patients and 63 from nonpatients. On a minute-to-minute basis, wrist activity alone was used to estimate Sleep Time. Blind independent scoring of the EEG-EOG-EMG records was also done to distinguish Sleep and Wake phases. Results from the two Sleep/Wake estimations agreed for 94.5% of the minutes (96.3% among nonpatients). Correlations between the two methods were determined for Total Sleep Period ($r = 0.90$), Total Sleep Time ($r = 0.89$), Wake After Sleep Onset ($r = 0.70$), and the number of Midsleep Awakenings ($r = 0.25$). Correlation coefficients were higher when the 39 patients were excluded from the computations. On the average, the actigraphic method overestimated Sleep Time by 15 min. Continuous wrist activity recordings provide simple and inexpensive, but rather accurate, estimates of sleep duration. Key Words: Sleep—Wake—Total sleep time—Sleep estimation—Wrist actigraph.

The two most common methods used to discriminate between the measures Sleep and Wake, behavioral observations and electroencephalographic (EEG) recording, both have major disadvantages. First, both methods are expensive. Much time is required to quantify sleep by scoring EEG recordings. Observational monitoring also requires large amounts of time and is nevertheless unreliable (Kupfer et al., 1970; Weiss et al., 1973), as is subjective estimation (Carskadon et al., 1976). Simpler and more cost-effective methods for quantifying sleeping and waking states are needed.

Activity recordings of various types have been used to distinguish Sleep and Wake. Sleep can be inferred in these recordings from the absence of activity. Thus, Levitt (1966) described an ultrasonic movement-sensing system which was used successfully to measure sleeping and waking behavior in small mammals. The minute-to-minute percentage agreement between EEG and activity measures of Sleep and Wake was 93%. Encouraging results with human subjects have been reported by Kupfer et al. (1972) and Foster et al. (1972*a*, 1972*b*), who employed

Accepted for publication August 1980.

Address correspondence and reprint requests to D. Mullaney, % Dr. Kripke at Department of Psychiatry (116), 3350 La Jolla Village Drive, San Diego, California 92161.

telemetric wrist-activity-recording devices. Their correlations between EEG and activity-based estimates of Sleep varied between 0.84 and 0.88. A more flexible wrist-activity-recording system was described by Kripke et al. (1978). In a pilot study with 5 subjects, Kripke et al. reported a 0.98 correlation between sleep duration estimated from wrist activity and the EEG.

In this report, we examine the reliability of this method with a much larger number of subjects, including hospital patients suffering from varying degrees and types of sleep disturbance.

METHODS

Subjects

Subjects included both males (n = 58) and females (n = 27) ranging in age from 18 to 66 ($\bar{X}$ age = 33.6, SD = 13.5). Sixteen subjects were recorded twice and 1, three times. Thirty-nine of the 102 recordings were obtained from hospitalized patients, selected to examine the effectiveness of the method for subjects suffering disturbed sleep. These patients had psychiatric problems, alcoholism, or chronic pain syndromes, and in addition to other complaints, almost all of them reported some degree of insomnia. The 63 recordings of nonpatients were obtained from a heterogeneous population including college students and hospital staff in good health who generally reported sleeping well.

Instruments and Procedure

The activity transducer or actigraph was constructed by soldering a small nut off center onto a 5 mm length of springlike wire and clamping the other end of the wire against a piezoceramic element (Kripke et al., 1978). Because the weight of the nut was off center, the piezoceramic element was excited when the transducer was moved in any direction, although the voltage produced was not linearly related to motion or acceleration. The transducer was packaged in a small acrylic box mounted on a watch band. A small Medilog tape recorder,[1] worn on a belt around the waist, recorded the transducer output.

In order to cross-validate the estimates of sleep from activity recordings, parietal-occipital EEG, submental electromyogram (EMG), and bipolar horizontal electro-oculogram (EOG) were recorded simultaneously with wrist activity. A preamplifier was used to amplify EEG and EMG so that these signals could be recorded with the other data onto a C-120 cassette. Four channels of information were thus recorded on the same cassette. A signal mark noting the time of going to bed was superimposed on the EMG.

In the afternoon, electrodes were attached to the subjects, and the activity transducer was mounted on the wrist. The dominant wrist was selected rather than the nondominant on the theory that more waking activity would be obtained. Subjects were instructed to go about their normal routines with the exception of any activities which could jeopardize the equipment. They were also asked to

[1] Ambulatory Monitoring, Inc., 731 Saw Mill River Rd., Ardsley, New York 10502.

press the signal marker upon entering bed. The following morning subjects filled out a sleep log in order to subjectively estimate sleep times and sleep quality. They slept at home or in their hospital beds (if they were inpatients) and returned to the lab at a convenient time the following morning for removal of the electrodes.

The recordings were played back at 60 times the recording speed and dubbed onto a ½ inch instrumentation tape along with a time code identifying each recorded minute. A segment of the total 16–18 hr recording was selected for analysis. The amount of Wake Time which surrounded the sleep period was varied from night to night between 1 and 5.5 hr in order to avoid cueing the scorers. EEG-EOG-EMG signals from this segment plus the time code were replayed onto polygraph paper to obtain a standard 15 mm/sec recording. Simultaneously, the analog activity recording for the same segment was replayed with the same time code onto a separate polygraph at 32 mm/min.

The EEG-EOG-EMG recordings were scored according to the criteria of Rechtschaffen and Kales (1968) in 1 min epochs. Scoring results were combined into two categories: Wake, which includes both Wake and Movement Time, and Sleep, including stages 1–4. A second scorer independently scored each minute of the actigraphic record for Wake or Sleep using the best guess which could be derived from activity alone. If greater than 50% of a 1 min epoch contained activity, this epoch was always scored as Wake. The context of the epoch was also considered, so that an epoch with less than 50% activity was scored Wake if it was surrounded by other epochs with sufficient activity. The activity scorer was blind to both the EEG records and sleep logs.

After about 60% of the records had been scored, it was realized that the technical quality of recordings (e.g., freedom from artifacts such as 60 Hz interference or EEG electrode loosening) and the interpretability of recordings would affect inter-rater agreement. Therefore, the later records were also rated as technically "high" or "low" in quality and their interpretability was likewise rated as high or low. Most of the difficult-to-interpret records were obtained from patients with sleep disorders. This prospective rating (before agreement was determined) was used to evaluate how well scoring unreliability could be recognized.

Since the validity of a new measure is limited by the reliability of the accepted measurement standard, 10 EEG records were scored independently by two raters so that EEG scoring reliability could be estimated. In order to estimate actigraph scoring reliability, duplicate activity recordings corresponding to the EEG records were also independently scored.

RESULTS

A total of 102 recordings from 53 nonpatients and 32 patients were obtained. The quality of the recordings using the Medilog system varied but was generally comparable to polygraphic EEG-EOG-EMG recordings. Technically inadequate or marginal recordings were excluded from analysis so that all 102 recordings used in this study were judged adequate for Rechtschaffen and Kales scoring. The most common technical difficulties causing records to be excluded were due to electrodes loosening, wires breaking, plugs disconnecting, tape jamming, cross-talk

between channels, and battery failures. The actigraphic transducer rarely failed. Problems were greatly reduced by use of fresh batteries and new cassettes. Among the records that were used and rated for quality, sources of poorer-quality recordings were bad EEG electrode impedances (4 cases), peculiar EEG patterns obtained from patients on medications (5), and one record with a bad EOG. Actigraph tracings rated "poor" contained noise due to "tape squeal" (9 cases), cross-talk from the EMG/clock channel (2), or minor damage to the equipment. All of these problems except the EEG peculiarities resulting from medications can easily be eliminated and the reliability of the records increased.

For each recording, the number of minutes for which the actigraph scorer and the EEG scorer agreed in assigning a designation of Wake or Sleep was determined, as well as the number of minutes where the two scorers disagreed. For the 102 recordings, the two scorers agreed on 94.5% of all minutes. The percent agreement for the 63 nonpatient recordings was 96.3%. For the 39 patient recordings, agreement was 91.6%.

A more meaningful measure than the percentage of agreement for evaluating the usefulness of the wrist actigraph for quantifying the Sleep measure is the extent to which the wrist actigraphic scoring reflects inter-subject and inter-night variability in Sleep Time. That is, can we differentiate short sleep, insomnia, etc., using this method? Across nights, the correlation coefficient between the two estimates for Total Sleep Time (TST) was $r = 0.90$ ($p < 0.0001$), indicating that the wrist-actigraphic scoring reflected most of the variance in EEG Sleep Time. As is shown in Table 1, the correlation coefficients between the two methods were also significant for estimating other sleep parameters—Total Sleep Period (TSP), i.e., the interval from the first to last Sleep, minutes of Wake After Sleep Onset (WASO), and the number of Midsleep Awakenings (MSA).

Although the minute-to-minute agreements and correlations were very high, we found that the actigraphic scorer's estimate of TST was, on the average, 15.33 min greater per night than the EEG estimate, and this difference was highly significant ($t = 3.82$, $p < 0.0001$). Examination of scoring errors showed that these errors were largely due to occasions when the subject was awake but did not move his or her wrist. This tendency of the actigraphic scorer to score Sleep when the EEG scorer indicated Wake is also demonstrated in Table 2. Table 2 shows that this type of disagreement was three times more frequent than disagreements where the actigraphic scoring was Wake and EEG scoring was Sleep. The average estimates of TSP, WASO, and MSA by the two methods were not significantly different.

Since the actigraphic method tended to overestimate TST by a mean of 15 min across all subjects, we counted the number of times the actigraph or the EEG rater scored sleep onset first. Out of 102 records, the actigraphic scorer scored sleep onset first in 56 instances, while the EEG scorer scored sleep onset first in 33 cases. In the remaining 13 sleep records, the two raters agreed to the minute when sleep onset occurred. The tendency for the actigraphic scorer to score sleep onset before the EEG scorer was highly significant ($p < 0.0001$), and the mean discrepancy was 4.18 min, which accounts for more than 25% of the discrepancy between the two methods.

To explore the relative reliability of the actigraphic method in various condi-

TABLE 1. *Correlation between actigraph and EEG scoring*

Group variable	TSP		TST		WASO		MSA		Agreement (%)	
All subjects (n = 102)	0.90[a]		0.89[a]		0.70[a]		0.25[c]		94.5	
Patients (n = 39)	0.82[a]	($p < 0.0001$)	0.81[a]	($p < 0.001$)	0.56[a]	($p < 0.001$)	0.09[e]	($p < 0.05$)	91.6	($p < 0.001$)
Nonpatients (n = 63)	0.97[a]		0.95[a]		0.87[a]		0.46[a]		96.3	
Age										
≥50 (n = 17)	0.52[a]	($p < 0.0001$)	0.39[e]	($p < 0.0001$)	0.82[a]	(ns)	0.01[e]	(ns)	88.5	($p < 0.001$)
<50 (n = 85)	0.95[a]		0.95[a]		0.65[a]		0.32[b]		95.7	
EEG TSP										
≥390 (n = 66)	0.85[a]	(ns)	0.83[a]	(ns)	0.72[a]	(ns)	0.19[e]	(ns)	94.7	(ns)
<390 (n = 36)	0.74[a]		0.78[a]		0.66[a]		0.52[a]		94.2	
Sleep log TST										
≥390 (n = 53)	0.96[a]	($p < 0.0001$)	0.94[a]	($p < 0.0001$)	0.72[a]	(ns)	0.26[d]	(ns)	96.3	($p < 0.001$)
<390 (n = 49)	0.79[a]		0.69[a]		0.70[a]		0.23[e]		92.5	
Recording quality										
Low (n = 20)	0.78[a]	($p < 0.001$)	0.79[a]	($p < 0.001$)	0.74[a]	(ns)	−0.07[e]	($p < 0.05$)	92.8	(ns)
High (n = 24)	0.97[a]		0.97[a]		0.75[a]		0.59[c]		95.9	
Interpretability										
Low (n = 27)	0.77[a]	($p < 0.001$)	0.83[a]	($p < 0.01$)	0.62[a]	(ns)	−0.16[e]	($p < 0.01$)	90.4	($p < 0.005$)
High (n = 24)	0.97[a]		0.97[a]		0.75[a]		0.59[c]		95.9	

[a] $p < 0.0001$ (one-tailed), [b] $p < 0.001$, [c] $p < 0.01$, [d] $p < 0.05$, [e] not significant.
Abbreviations: TSP, total sleep period; TST, total sleep time; WASO, wake after sleep onset; and MSA, number of midsleep awakenings.

TABLE 2. *Comparison of EEG and actigraphic scoring*[a]

	EEG scoring (min)	
	Sleep	Wake
Actigraphic scoring (min)		
Sleep	378	24
Wake	8	172

[a] Average for 102 recordings.

tions, we split out samples of 102 nights in six different ways, namely, patients vs nonpatients, subjects 50 and older vs. subjects younger than 50, subjectively short or long sleep (sleep log less or more than 390 min), objectively short or long sleep (EEG sleep less or more than 390 min), quality of recording (low or high), and interpretability (low or high). In each case, we predicted the former condition would produce more difficult scoring. Transformations of r to z were done so that a parametric test described by Hays (1973) could be used to determine if the correlation coefficients for dichotomous categories were significantly different (Table 1). Significantly higher correlations for nonpatients were obtained for all four sleep parameters. The differences were also highly significant in the predicted directions for TST and TSP in the dichotomies of age, sleep log TST, and quality and interpretation ease, but correlations for subjects whose EEG TSP was >390 min were not significantly better. Comparable statistical results were obtained when t-tests were used to determine if percent agreement values in each category were significantly different (Table 1). In general, the patients tended to have short sleep log TSTs and poor interpretability and more were over 50 years old; so these characteristics were all associated.

It is notable that much poorer correlations were obtained among records of poorer technical quality—records rated as lower in quality in either the EEG-EOG-EMG or actigraphic recording. These reduced the overall averages.

The last 10 recordings both of EEG-EOG-EMG and of wrist activity were scored independently by both raters so that inter-rater percent agreement and reliability correlations could be computed for the EEG-EOG-EMG scoring and for activity scoring (Table 3). The inter-rater reliability correlations for estimating all sleep parameters were extremely high for both types of records, and although 7 of the 10 records were from patients, it happened that the EEG–actigraph correlations for these 10 records were higher than for the whole group of 102 records except for MSA, which had correlated poorly in other records from patients. For TST and WASO, the EEG–actigraph correlations were almost as good as inter-rater EEG correlations. These excellent results were obtained where all 10 EEG traces and 8 of 10 actigraphic records were rated high in quality. Once again, between-method reliability can be elevated by easily achieved improvements in technical recording quality. Table 3 also suggests that a substantial portion of EEG–actigraph scoring discrepancies were due to misscoring of the EEG records, despite the fact that our EEG inter-rater agreements were high.

TABLE 3. *Reliability estimates of EEG and actigraphic (ACT) scoring*

Comparison	Correlation coefficient (n = 10)				Agreement (n = 10) (%)
	TSP	TST	WASO	MSA	
EEG–EEG	0.999 ($p < 0.0001$)	0.977 ($p < 0.0001$)	0.968 ($p < 0.0001$)	0.899 ($p < 0.0001$)	96.5
ACT–ACT	0.981 ($p < 0.0001$)	0.979 ($p < 0.0001$)	0.988 ($p < 0.0001$)	0.851 ($p < 0.001$)	97.2
EEG–ACT	0.976 ($p < 0.0001$)	0.956 ($p < 0.0001$)	0.964 ($p < 0.0001$)	−0.079 (ns)	92.0

Abbreviations same as in Table 1.

Correlation coefficients were also computed between sleep log TST and the two objective TST estimates in order to determine how the actigraphic and EEG estimations of Sleep compared to subjective estimations. The correlation of EEG and sleep log TSTs was $r = 0.83$ ($p < 0.0001$), while the correlation of actigraph and sleep log TSTs was somewhat lower ($r = 0.75$, $p < 0.0001$). These two correlations did not differ significantly. As mentioned earlier, the correlation between the two objective methods of estimating Sleep was $r = 0.90$ ($p < 0.0001$). Thus, as referenced to EEG, the actigraphic method had less than 60% the error of subjective estimates, despite the fact that subjective estimates in this sample were rather reliable.

DISCUSSION

Our present results for estimating TSP, TST, and WASO in nonpatients ($n = 63$) of $r = 0.97$, 0.95, and 0.87, respectively, are comparable to our pilot results ($n = 5$), where we obtained correlation coefficients of 0.95, 0.98, and 0.85, respectively (Kripke et al., 1978). Our results with patients considered separately were not so satisfactory, and the actigraph failed to reliably note MSA in the patient group. Nevertheless, the results summarized in Table 3 suggest that except for MSA, EEG inter-rater agreements are not remarkably better than EEG–actigraph agreements in a predominantly patient group. Indeed, our minute-to-minute actigraph–EEG agreement for scoring Wake is actually better than the EEG–EEG agreement recently reported in a careful study of normal records (Karacan et al., 1978). It is also clear that much of the difficulty scoring actigraph records was due to technically poor recordings or to problems of interpretation which could be recognized prospectively. Few of the problems, however, were due to the actigraphic method, and many problems could be eliminated by use of the newer-model Medilog recorders (ours were among the first built), new tapes, new batteries, and occasional repeat recordings. It is not entirely clear that the actigraphic method is relatively deficient with records of patients and older subjects, because there is inadequate evidence that EEG scoring is reliable in such groups. Table 3 shows that many discrepancies between actigraph and EEG may be attributed to difficulties with EEG scoring. The older patients with disturbed sleep, especially those with sleep apnea and nocturnal myoclonus, arouse so frequently that the polygraphic patterns sometimes lose their resemblance to the Rechtschaffen and Kales model (1968). In these altered states of physiologic function, our usual classifications of wakefulness and sleep become blurred, and highly reliable scoring is no longer possible. Further studies are needed to determine if EEG scoring of patient records is actually much more reliable than that obtained by actigraph.

In future studies employing actigraphic scoring to estimate TST, it will be important to bear in mind the significant tendency of the actigraph to overestimate TST in comparison to EEG estimation. Scoring in the more traditional 20 or 30 sec epochs may reduce this actigraphic overestimation of TST, since we scored 1 min epochs as Sleep when less than 30 sec contained activity. Shorter epochs should lead to greater scoring precision. Even though the mean increase of 15 min is a small amount of time—only about 3% of an 8 hr sleep period—the actigraphic overestimation may be much greater with some individuals than with others,

depending on idiosyncratic behavior patterns. For example, one subject "fooled" the actigraph scorer after awakening by sitting perfectly still and meditating for 20 min before breakfast.

In this study, we have demonstrated that the wrist actigraph is a valid and reliable instrument for quantifying sleep time as compared to standard EEG-EOG-EMG scoring, especially when ambulatory recording is necessary. Most investigators would probably assume that EEG scoring corresponds more accurately to behavioral and functional sleep than does actigraph scoring and, therefore, that EEG scoring is more accurate. The subjective data tended to support this assumption. While we have no objective comparisons of EEG and actigraphic scoring as indicators of behavioral sleep, our reliability data indicated that the agreement between EEG and actigraphic estimates of TST is only slightly less than agreement of two scorers rating the same EEG record. Therefore, the reliability of the wrist actigraph must be rather close to EEG accuracy for quantifying behavioral and functional sleep.

Even if the EEG method is slightly more accurate in estimating sleep time, a cost–benefit analysis would often favor the actigraphic method. Increased accuracy in a sample measurement is only valuable to the extent that the sample precisely represents the population from which it has been drawn. Night-to-night and subject-to-subject variability in TST is so substantial that very high accuracy is usually wasted. For example, among the 17 subjects for whom 2 nights were recorded, our night-to-night correlations for EEG TST were only $r = 0.392$. Other studies have reported correlations ranging from $r = 0.127$ (Moses et al., 1972) to $r = 0.50$ and 0.78 (Johnson et al., 1970). Perhaps the most thorough study of TST correlations between 4 nights reported correlations from $r = 0.03$ to 0.68 (Clausen et al., 1974). In our study the mean night-to-night (within-subjects) standard deviation of EEG TST was 59.18 min, while the standard deviation of EEG–actigraph discrepancies for identical nights was 23.95 min. Thus, for example, if we obtained 4 nights of EEG recording for each subject, the mean standard error of the mean TST for each subject would be 34.17 min. On the other hand, the combined effect of night-to-night variability and method-to-method discrepancy can be estimated by adding the night-to-night EEG sleep-time variance to the EEG–actigraph variance. We find that the standard error of the mean actigraphic measures (including the discrepancies with the EEG estimates) would be 36.9 min for 4 nights of recording. Let us further suppose that the cost of actigraphic recording were only 1/5 that of EEG recording (we believe it may actually be closer to 1/10). Then, for an equal investment, we could obtain 20 nights of actigraphic recording, leading to a standard error of the mean TST of 14.27 min. Thus, for an equal cost, the actigraphic method would be at least twice as precise as EEG for estimating mean TST. Other factors make the actigraph a more desirable method for many research and clinical purposes, studies of industrial health, and clinical screening. The actigraphic method produces less skin irritation and has a lower shock risk. The actigraph is also more comfortable and convenient for the subject, who can sleep in his own bed if desired, free from electrodes. Also, the adaptability of the actigraphic method to naturalistic settings improves its verisimilitude as a measurement technique.

Thus, the actigraphic method of measuring sleep is a reliable method which actually allows more precise measurement of EEG TST than the EEG recording method for a given cost. Nevertheless, there is room for improvement because the actigraphic method may benefit from computer technology. We believe that an all-digital actigraphic recorder can be developed which can eliminate the need for the Medilog recorder, and which will ultimately be adaptable to the wrist. We are now testing a computer program which makes scoring completely automatic. Technologic development towards this goal, which should provide a still greater improvement in the cost–benefits ratio of the method, is the focus of our current work.

ACKNOWLEDGMENT

This work was supported in part by DAMD 17-78-C-8040, by the Medical Research Service of the Veterans Administration, by ONR N00014-79-C-0317 and by Research Scientist Development Award 1 K02 MH00117 (to D.F.K.) from the National Institute of Mental Health.

REFERENCES

Carskadon MA, Dement WC, Mitler MM, Guilleminault C, Zarcone VP, and Spiegle R. Self-reports versus sleep laboratory findings in 122 drug-free subjects with complaints of chronic insomnia. *Am J Psychiatry* 133:1382–1388, 1976.

Clausen J, Sersen EA, and Lidsky A. Variability of sleep measures in normal subjects. *Psychophysiology* 11:509–516, 1974.

Foster FG, Kupfer D, and McPartland R. Telemetric mobility recording in assessment of day and night activity cycles. In: MH Chase, WC Stern and PL Walter (Eds), *Sleep Research*, Vol 1, Brain Information Service/Brain Research Institute, UCLA, Los Angeles, p. 171, 1972*a*.

Foster FG, Kupfer D, Weiss G, Lipponen V, McPartland R, and Delgado J. Mobility recording and cycle research in neuropsychiatry. *J Indiscipl Cycle Res* 3(1):61–72, 1972*b*.

Hays WL. *Statistics for the Social Sciences*, 2nd ed. Holt, Rinehart & Winston, New York, 1973, pp 663–664.

Johnson LC, Burdick JA, and Smith J. Sleep during alcohol intake and withdrawal in the chronic alcoholic. *Arch Gen Psychiatry* 22:406–418, 1970.

Karacan I, Orr, WC, Roth T, Kramer M, Shurley JT, Thornby JI, Bingham SF, and Salis PJ. Establishment and implementation of standardized sleep laboratory data collection and scoring procedures. *Psychophysiology* 15:173–180, 1978.

Kripke DF, Mullaney DJ, Messin S, and Wyborney VG. Wrist actigraphic measures of sleep and rhythms. *Electroencephalogr Clin Neurophysiol* 44:674–676, 1978.

Kupfer DJ, Wyatt RJ, and Snyder F. Comparison between electroencephalographic and systematic nursing observations of sleep in psychiatric patients. *J Nerv Ment Dis* 151:361–368, 1970.

Kupfer DJ, Detre TP, Foster F, Tucker GJ, and Delgado J. The application of Delgado's telemetric mobility recorder for human sleep studies. *Behav Biol* 7:585–590, 1972.

Levitt RA. An activity measure of sleeping and waking behavior. *Psychon Sci* 5:287–288, 1966.

Moses J, Lubin A, Naitoh P, and Johnson LC. Reliability of sleep measures. *Psychophysiology* 9:78–82, 1972.

Rechtschaffen A and Kales A. *Manual of Standard Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects*. Brain Information Service/Brain Research Institute, UCLA, Los Angeles, 1968.

Weiss BL, McPartland R, and Kupfer DJ. Once more: The inaccuracy of non-EEG estimates of sleep. *Am J Psychiatry* 130:1282–1285, 1973.