

SCIENTIFIC INVESTIGATIONS

Examining Initial Sleep Onset in Primary Insomnia: A Case-Control Study Using 4-Second Epochs

Douglas E. Moul, M.D.; Anne Germain, Ph.D.; J. David Cashmere, B.S.; Michael Quigley, B.S.; Jean M. Miewald, B.A.; Daniel J. Buysse, M.D.

University of Pittsburgh Department of Psychiatry, Pittsburgh, PA

Study Objectives: To explore the sleep onset process in primary insomnia patients, new rules for scoring 4-second epochs were implemented to score sleep and artifacts during initial sleep onset. Conventional scorings in 20-second and 60-second epochs were also obtained.

Methods: The start of the initial 60-second epoch of stage 1 was used to define "time zero" (t_0). Sleep onset periods from 11 patients and 11 individually age- and sex-matched controls spanned from 5 minutes before t_0 through 29 minutes after t_0 . Using the new rules, the periods were scored blind to group assignment. This t_0 time-referenced the data analysis to one plausible midpoint in the sleep onset process. In parallel, latencies were time-referenced from good night time.

Results: Reliability in scoring sleep and artifacts was adequate ($\kappa = 0.68$ & 0.63 , respectively, $p < 0.001$). Group differences in sleep latencies were marginal in 60-second and 20-second scoring but significant with a definition of 4-second sleep latency. Patients had more 4-second

epochs scored as awake (Mantel-Haenszel $\chi^2 = 271$, d.f. = 1, $p < 0.001$) and containing artifact (M-H $\chi^2 = 143$, $p < 0.001$). Patients took longer to achieve 30 continuous 4-second epochs of NREM sleep (Breslow $\chi^2 = 4.03$, d.f. = 1, $p = 0.045$) after t_0 . Patients accumulated sleep more slowly with all 3 scoring rules after t_0 . A slower rate of accumulating sleep after t_0 was detected only with the 4-second scoring ($p = 0.047$).

Conclusions: Evidence was present for momentary state-switching instabilities in the patients during the initial sleep onset process. Using rules for scoring small epochs may reveal such instabilities more readily than traditional scoring methods.

Keywords: Insomnia, sleep onset, sleep latency, polysomnography, sleep scoring

Citation: Moul DE; Germain A; Cashmere D; Quigley M; Miewald JM; Buysse DJ. Examining initial sleep onset in primary insomnia: a case-control study using 4-second epochs. *J Clin Sleep Med* 2007;3(5):479-488.

INTRODUCTION

Sleep is scored as a state¹ but is a time series process. Scoring sleep involves classifying epochs of 20 seconds or 30 seconds (and formerly 60 seconds) as noncomplex states. These epoch lengths are longer than many arousals occurring in the process of sleep. Arousals and artifact summaries are not part of traditional sleep architecture descriptions.

While some primary insomnia patients have abnormalities in conventionally defined sleep architecture,² for others the differences can be mild.^{3,4} Indeed, some noncomplaining poor sleepers have poorer architectures than some patients.⁵ This might suggest that the main pathology in patients might concern sleep/wake state misperceptions. Against this, studies point to increased beta power in the sleep EEG of insomnia patients during NREM sleep,⁶⁻¹⁰ suggest-

ing sleep process abnormalities. Analyzing EEG sleep processes exhaustively would implausibly require a wave-by-wave analysis. Small-epoch scoring may be an alternative.^{11,12}

Small-epoch scoring has a potential importance for insomnia research. Saper et al¹³ have proposed that defects in sleep/wake state-switching circuits within the brain may favor sleep fragmentation. In insomnia patients, sleep/wake states may be more variable¹⁴ or atypical,¹⁵ and the transition between states may be shallower and less stable.¹⁶ Sleep fragmentation would appear as small epochs of wake within larger, conventionally sized epochs.

Instabilities in sleep/wake state switching may be more likely to be observed during the sleep onset process. The sleep onset process can be considered as extending from full wakefulness through well-established stage 2 sleep.¹⁷ Figure 1 displays a schematic contrast of potential consequences of a defective state-switching system during the sleep onset process in insomnia patients. Helping patients get to sleep, or return to sleep, is often a key clinical task.¹⁸ Being unable to get to sleep initially can be the sole chief complaint.

The literature concerning the sleep onset process in insomnia^{6,10,15,16,19-23} includes studies that vary not only in sleep/wake state definitions, but also in time referencing (i.e., what is the time zero reference for the data?), sleep latency definitions (most elaborately addressed by Bonato¹⁶ who compared 7 definitions), and time interval (by quartile, fixed-length epoch, etc.) stipulations. This methodological variation is inevitable when one is attempting to describe a process that has vague boundaries and no obvious internal clock. Hori and colleagues have examined the sleep onset process in normal sleepers using detailed rules for scoring 5-sec-

Disclosure Statement

This is not an industry supported study. Dr. Buysse has been a consultant for Acetion, Cephalon, Eli Lilly, GlaxoSmithKline, Merck, Neurocrine, Neurogen, Pfizer, Respiroics, Sanofi-Aventis, Servier, Sepracor, and Takeda. Dr. Moul, Dr. Germain, Mr. Cashmere, Mr. Quigley, and Ms. Miewald have indicated no financial conflicts of interest.

Submitted for publication October, 2006

Accepted for publication March, 2007

Address correspondence to: Douglas E. Moul, M.D., M.P.H., Western Psychiatric Institute and Clinic, E-1119, 3811 O'Hara Street, Pittsburgh, PA 15213; Tel: (412) 246-5040. Fax: (412) 246-5300. Email: moulde@upmc.edu.

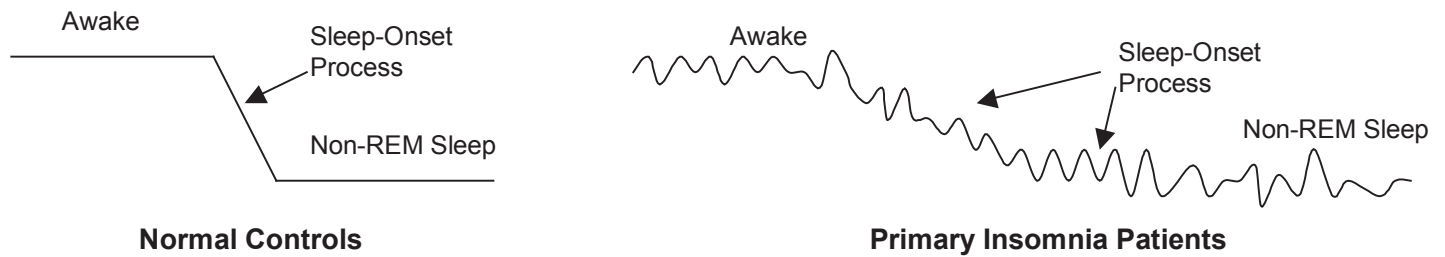


Figure 1—Conceptualization of possible time-series differences between Normal Controls and Primary Insomnia Patients: During Wakefulness, Non-REM Sleep, and the Sleep Onset Process. Heuristic model for possible abnormalities in sleep/wake processes in primary insomnia. Insomnia patients may exhibit greater fluctuations during sleep/wake states, or exhibit more sluggish and irregular state transitions.

ond epochs, classified into 9 states specifically related to the sleep onset process.^{24,25} Doerfling and colleagues applied the Hori rules in insomnia patients and controls using 7-second epochs, dividing the sleep onset process into quartiles.²⁴ They were unable to discriminate between groups with this method but did not analyze sleep latency or fixed-length epoch data.

At present, the Hori states do not have direct thalamocortical correlates. By contrast, wake, NREM, and REM sleep do. Perhaps scoring sleep in small epochs in these 3 simple states may have advantages. Compared to conventional rules, small-epoch scoring rules may increase statistical power, help identify state switching abnormalities, and assist in identifying insomnia subtypes. At present, conventional polysomnography of insomnia patients has little clinical utility,²⁶ so small-epoch scoring might have future utility.

It is often taken as obvious that EEG artifacts should play no role in scoring sleep. Artifact scoring is not part of current sleep staging. However, movement arousal artifacts may be abnormalities in the sleep of some insomnia sufferers.^{27,28} Artifact scoring may have value when scoring insomnia patients' sleep or sleep onset processes, and so is worth investigating.

This exploratory study compares the sleep onset process of primary insomnia patients and individually-matched controls. It utilizes a new method of classifying 4-second epochs into 3 sleep/wake states and identifying whether they contain one or more artifacts. The first aim of this study was to describe potential sleep onset process abnormalities in insomnia patients. A second aim was to explore how sleep latency metrics may correlate with self-reports in patients.

METHODS

Subjects

Eleven primary insomnia subjects (5 males, 6 females) who had participated in prior research protocols were identified from archival records of the Clinical Neuroscience Research Center at the University of Pittsburgh. Primary insomnia had been diagnosed in accordance with DSM-IV criteria²⁹ after screening with the Structured Clinical Interview for DSM-IV³⁰ and clinical interview, as well as physical and laboratory examinations to rule out other medical or sleep disorders.³¹ Each patient was individually age (± 3 years) and sex-matched to a control subject from the same archives, who had been assessed with identical procedures. All subjects had signed informed IRB-approved consent. All had been found free of clinically significant sleep disordered breathing or periodic limb movement arousals as assessed with oximetry,

nasal thermistor, piezoelectric belt, chin EMG, pretibial EMG, C_4 - A_1A_2 , and EOG channels. Apnea-hypopnea and periodic limb movement indices were less than 15 in all subjects. No subject was taking a medication for sleep at the time of study.

Definition of Scoring Rules for Three Epoch Lengths

The archival epoch scoring rules for 60-second epochs that time references this analysis were virtually identical to Rechtschaffen and Kales criteria² and had been used in several seminal publications from our laboratory. The visual scorings were maintained with good reliability ($\kappa = 0.74$) for the years of the study data. Sleep latency according to the local 60-second scoring rules was defined as the time length from the technician-identified time point when the subject attempted to initiate sleep (“good night time”) until the first stage 2 epoch in a 10-minute period of stable sleep, permitting 1 minute of wake or two minutes of stage 1 within that period.

The records were rescored for ASDA arousals³² and long arousals by a technician blind to group membership. ASDA arousals are defined as arousals lasting more than 3 seconds in a 30-second epoch of sleep. “Long arousals” were defined as waking events longer than 30 seconds. Both arousal indices were summarized in units of events per hour of sleep across the study nights. ASDA arousals were scored for their clock times and durations.

The records were blindly rescored by Rechtschaffen and Kales criteria in 20-second epochs. Recent percent agreement estimates from our laboratory were in the range of 0.85-0.95 for stage 2 sleep, but in the range of 0.50-0.85 for stage 1 sleep. The sleep latency definition for 20-second scoring was defined as the time from good night time until the first stage 2, 3, 4, or REM epoch, permitting not more than 1 consecutive minute of awake, or not more than 2 consecutive minutes of stage 1, or not more than 2 minutes of any combination of awake or stage 1 in the following 10 minutes.

Study rules for visually scoring 4-second epochs were constructed by a consensus of the authors (see Appendix). The rules classified epochs by 1) wake, NREM, or REM state and 2) artifact presence. Artifact presence classification was guided by a priority for specificity in the definition of epochs free of muscle, EKG, eye movement, electrode, and other artifacts. Any artifact present, including eye movements, denoted an epoch as having an artifact. Artifacts were not subtyped. The sleep latency for the 4-second scoring rules was defined as the period from good night time until the first epoch of a 10-minute period in which 90% of epochs were classified as NREM.

Data Acquisition and Processing

Each subject's record was selected randomly, provided it was not the adaptation night. Each record had originally been scored in 60-second epochs. The oldest was acquired in 1994, the newest in 2000. Standard post-sleep self-report data for the nights studied were also obtained.

The first 60-second epoch of stage 1 was identified from the archival records for the main time alignment for the data analysis. Its start defined "time zero" (t_0). This choice was made in order to identify a time-reference point within the sleep onset process itself as originally scored. Based on this t_0 reference, 5 prior minutes and 29 subsequent minutes of C_4 - A_1 - A_2 recording were extracted from the archives. In some cases, t_0 occurred earlier than 5 minutes from the record's beginning. In these cases "wake" was imputed for these prior epochs. The first author, who was blind to group membership of subjects, scored all records with the new rules for 4-second epochs. The second author scored 4 of the records in the same manner to provide data concerning the approximate reliability of the rules. One patient had REM sleep start late in the record. To confine the analysis only to NREM, all records were shortened to 33 minutes, comprising the 495 four-second epochs that were analyzed.

Each subject's progressive accumulation of sleep across the 4-second epoch series was computed. For the 4-second epoch data this was done directly. For the 20-second and 60-second epoch data, each 4-second epoch was assigned the sleep/wake state (Wake versus NREM) of the longer epoch of which it was a member. For mixed effects analyses between t_0 and 60 seconds afterward, the individual cumulative 4-second and 20-second epoch scoring data were analyzed in their as-scored form.

Other Time Alignments

In this analysis, "latency" refers to the time interval until a defined event as measured from the good night time starting point. To test time period relationships within the sleep onset process, the time intervals between t_0 and the sleep latency points as separately defined under the 60-second, 20-second, and 4-second versions of sleep latency end points (see above) were calculated.

To create models of sleep continuity after t_0 , different models of sleep continuity in 4-second epochs were constructed. These models were based on the criteria of unbroken continuity in a 4-second epoch series. The tested series lengths were for 8, 15, and 30 continuous 4-second epochs (i.e., 32 seconds, 60 seconds, and 120 seconds). The time point after t_0 that ended the first unbroken series of 4-second epochs defined the endpoint in these time-to-event analyses.

To test for group differences in timing of ASDA arousals after t_0 , intervals from t_0 to the first arousal were computed for Kaplan-Meier analysis. The intervals between the first and second, and second and third arousals were also computed.

Statistical Analysis

A significance level of $p = 0.05$ was set for all analyses. The 20-second sleep latency definition was taken as the sleep latency best representing conventional scoring practices. Since the analysis was focused on elaborating a process description, a multimethod approach was necessary. Statistical tests can evaluate processes

by finite data that are referenced at time points, and not by non-finite data that are themselves processes. To compensate for this limitation, multiple endpoints and frames of reference needed to be evaluated for the process description to be elaborated. Using just one endpoint such as sleep latency would be insufficient. The study's principal aim was to bring out sleep onset process abnormalities in comparison to conventional methods of scoring sleep latency. Separate evaluations of conventional scoring results, reliability testing of the new scoring rules, analyses of event frequencies, tests of time-to-event models, and regression models of cumulative sleep epochs supported meeting this aim.

Paired t -testing was performed on sleep parameter values and self-report data, matching each patient to his/her control. Reliability testing was done using the kappa statistic in SPSS,³³ which adjusts for the expected base-rate agreement. Mantel-Haenszel χ^2 tests stratified by patient:control matches were used to test for group differences in probabilities of 4-second epochs with sleep or artifacts. The Wilcoxon rank test was used to test for differences between groups in the distribution of the ASDA arousal event times and the event times of epochs containing artifacts after t_0 . Kaplan-Meier life-table analysis using chi-square testing using the Breslow method in SPSS³³ was used to test for group differences in time-till-event analyses. For the study of the beginning time of sleep continuity as judged with 4-second epoch time series, sleep continuity was modeled as series of 8, 15, and 30 continuous four-second epochs of NREM. For time-till-event analyses of ASDA arousals, Kaplan-Meier analyses were conducted for the intervals between t_0 and the first arousal, between the first and second arousal, and between the second and third arousal.

Linear regression models of the cumulative distributions of 4-second epochs were constructed with S-PLUS³⁴ software to model time-trend data from all 3 scoring rules. In these models, the dependent variable was the cumulative mean sleep in each group from t_0 through all the subsequent 420 epochs. Stepwise regressions added group and group-by-time terms. Analysis of variance tests determined if adding a parameter improved model fit. Such models addressed the general time scale of conventional sleep latencies.

To address the more moment-to-moment perspective of the sleep onset process, linear mixed effects models using maximum likelihood estimation were constructed where the dependent variable was the cumulative sleep in 4-second scoring data at t_0 and at 20 seconds, 32 seconds, and 60 seconds, respectively. The interval between t_0 and these endpoints approximated 20-second, 30-second, and 60-second epoch lengths. Starting models included fixed- and random-effect terms for time, blocked by individual. Stepwise models added group (fixed term) and group-by-time (both fixed and random) interaction terms. Likelihood ratios between models were tested with analysis of variance to see if the added terms significantly improved the model fit. If the fixed interaction term improved the model, then the groups (fixed term) differed in the rate of sleep build-up, after adjusting for individual differences. Similar models tested the 20-second scoring data, to test whether using it could similarly detect such momentary group differences in sleep build-up.

For the study's second aim, Pearson correlations tested relationships between post-sleep self-report variables versus t_0 latency, the 3 sleep latencies, and the t_0 to 20-second sleep latency interval.

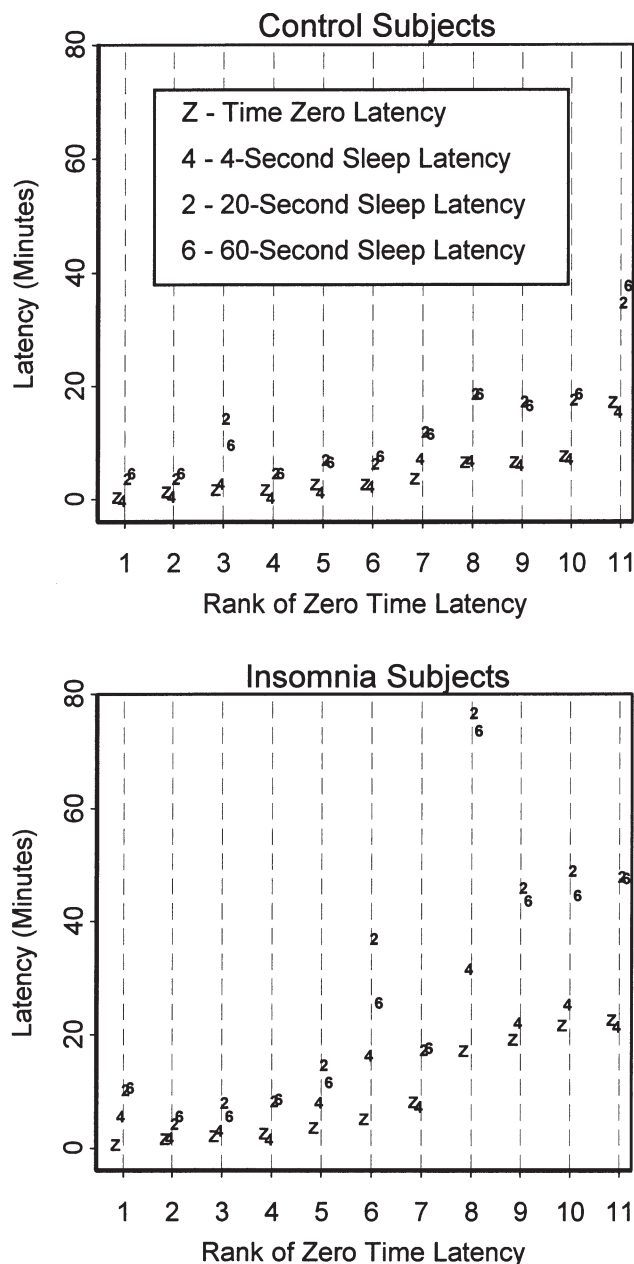


Figure 2—Graph of Latencies. Latencies from technician-defined starting points for attempting to sleep until defined time points are shown, ranked in order of increasing latencies to the t_0 reference point. Time points are coded: Z – Time Zero (beginning of first 60-second epochs scored as Stage 1); 4 – Sleep latency according to the stipulated definition for 4-second scoring; 2 – Sleep latency according to the stipulated definition for 20-second scoring; 6 – Sleep latency according to the stipulated definition for 60-second scoring. The insomnia subject # 1 was ranked 8th in this graph, for which the 4-second sleep latency was imputed.

RESULTS

Sample Description

The patients (5 male, 6 female; age range 24– 57; mean 45.3 years \pm 11.3 SD) were drawn from 3 separate protocols. They reported a Pittsburgh Sleep Quality Index³⁵ (PSQI) score of 10.7 ± 3.4 . On the day prior to the study night, 3 had consumed one caffeinated beverage and 8 had consumed no caffeinated beverages. None had taken a nap nor consumed alcohol.

The controls (5 male, 6 female, age range 24–57; mean 44.8 years \pm 10.7) had PSQI scores of 1.4 ± 1.12 . All were drawn from a single protocol. On the day prior to the study night, 1 had consumed four caffeinated beverages, 2 had consumed two, 3 consumed one, and 5 consumed no caffeinated beverage(s). None had taken naps or consumed alcohol.

Reliability of the 4-Second Epoch Scoring Rules

Reliability of the scoring rules was good both for sleep state (% agreement = 89%, kappa = 0.68, $p < 0.001$), and for classification of epochs with artifacts (% agreement = 90%, kappa = 0.63, $p < 0.001$).

Conventional Sleep Stage Scoring and Latency Comparisons

Conventional polysomnographic sleep parameters are presented in Table 1, both for 60-second and 20-second scoring rules. There were no statistically significant differences in total sleep time or for the ASDA arousal or long arousal indices between patients and controls.

According to the $p = 0.05$ criterion with paired t -testing, there was only a trend in group differences in sleep latency for 60-second scoring, but there was a significant difference for the 20-second scoring definition, reflecting effect sizes of 0.64 and 0.68, respectively, in the paired analyses. In the Kaplan-Meier analysis of the 60-second and 20-second data, there were no group differences in these sleep latencies.

There was a trend difference between groups in latency to t_0 by t -testing ($t = 1.9$, $df = 10$, $p = 0.09$, effect size = 0.57), but not by Kaplan-Meier analysis. With t -testing, the groups did not differ in the interval between t_0 and the respective 60-second and 20-second sleep latency time points ($t = -1.31$, $df = 10$, $p = 0.2$; $t = -1.42$, $df = 10$, $p = 0.2$, respectively). There was only a trend difference ($p = 0.09$) in Kaplan-Meier analysis of t_0 to 60-second sleep latency.

As shown in Figure 2, the sleep latency point according to the 4-second definition approximated t_0 in most subjects. A key exception was insomnia match subject # 1, who did not attain 4-second sleep latency within the range of the data. This subject's 4-second sleep latency was imputed to the end of the data. Inspection of Figure 2 suggested a linear proportionality between t_0 latency and the 20-second scoring rule latency. These correlated 0.95 ($p < 0.001$) and 0.84 ($p = 0.001$) for the controls and patients, respectively. Notably, there were group differences in 4-second sleep latencies ($\chi^2 = 4.3$, $df = 1$, $p = 0.04$), present even without the first match pair included ($\chi^2 = 4.6$, $df = 1$, $p = 0.03$). Some 4-second sleep latencies preceded t_0 , making Kaplan-Meier analysis of the t_0 to 4-second sleep latency interval invalid. With t -testing, the patients were found to have more 4-second sleep latencies after t_0 ($t = 2.43$, $df = 10$, $p = 0.036$); with the trend persisting even with first match pair removed ($t = 2.206$, $df = 9$, $p = 0.055$). No group differences were found with Kaplan-Meier analysis in the interval between the 4-second latency point and the 60-second or 20-second sleep latency points.

The insomnia patients did not have a statistically significant difference (see Table 2) in their self-estimation of sleep latency on the study night compared to their matched controls, even while reporting that their sleep latency was usually longer ($p = 0.015$). They tended to self-report shorter total sleep time both for the

Table 1—Standard Polysomnographic Parameters

Item	Control mean (SD)	Insomnia mean (SD)	t (df = 10) ¹	p
Clock Time of Bedtime	23:08 (48)	23:09 (72)	0.050	0.961
Clock Time Out of Bed	06:27 (42)	06:54 (85)	0.985	0.348
60-Second Epoch Scoring				
Sleep Latency (minutes)	13.2 (9.9)	27.2 (22.4)	2.105	0.062
Total Spent Asleep (minutes)	391.0 (34)	384.0 (47)	-0.487	0.637
Time Awake after Sleep Onset (minutes)	35.0 (32)	54.0 (45)	1.154	0.275
Awakenings (number)	6.3 (3.7)	7.5 (4.0)	0.892	0.394
Stage 1 (%)	4.7 (2.4)	6.5 (3.3)	2.114	0.061
Stage 2 (%)	61.9 (8.3)	57.2 (10.1)	-1.063	0.313
Delta (%)	8.8 (10.0)	11.8 (10.6)	0.618	0.550
REM (%)	34.6 (2.7)	34.5 (5.5)	-0.005	0.996
ASDA arousals per Hour	7.8 (2.0)	8.9 (2.2)	0.891	0.394
Arousals per Hour	1.0 (0.6)	1.2 (0.7)	0.740	0.477
Sleep Efficiency (%)	89.1 (6.9)	82.7 (8.1)	-1.733	0.114
Sleep Maintenance (%)	91.8 (7.2)	87.9 (9.6)	-1.158	0.274
20-Second Epoch Scoring				
Sleep Latency (minutes)	13.1 (9.3)	29.4 (23.5)	2.263	0.047
Total Spent Asleep (minutes)	386.0 (34)	342.0 (50)	-0.995	0.343
Time Awake after Sleep Onset (minutes)	40.0 (31)	64.0 (50)	1.415	0.187
Awakenings (number)	17.7 (3.4)	19.2 (6.1)	0.720	0.488
Stage 1 (%)	7.1 (5.3)	5.8 (2.8)	-0.745	0.474
Stage 2 (%)	62.3 (6.0)	58.6 (8.2)	-1.341	0.209
Delta (%)	7.2 (6.4)	10.4 (8.4)	1.154	0.275
REM (%)	23.4 (3.9)	25.2 (5.0)	0.949	0.364
Sleep Efficiency (%)	88.0 (6.6)	80.3 (9.4)	-2.049	0.068
Sleep Maintenance (%)	90.7 (7.0)	85.5 (10.7)	-1.478	0.170

¹Paired *t*-tests on matched control-patient pairs.

study night and generally more wakefulness than controls. They felt a modest degree of insufficient sleep, more poorly rested, and more confused and depressed than the controls. They reported greater difficulty awakening.

Categorical Analyses of 4-Second Epoch Data

Time series of the different scoring epoch data are shown in Figure 3. Upon inspection, the archival 60-second scoring for insomnia patient #1 (See Figure 3, Part A) appeared in disagreement from the 20-second (Figure 3, Part B) and 4-second scorings (Figure 3, Part C). A technician rescored this one record blindly in 60-second epochs (see isolated scoring in between control and insomnia blocks, Figure 3, Part 1). Stage 1 was rescored to be a minute earlier than in the archive, and many of the subsequent epochs were awake. But this rescoreing was in general agreement with the intended t_0 as a rough process midpoint. The original scoring was kept as the basis for t_0 in this record, since the record's temporal orientation was faithful to the kind of results obtained from typical scoring procedures.

Compared to controls, patients had more 60-second epochs scored as wake (Mantel-Haenszel $\chi^2 = 5.8$, $df = 1$, $p = 0.016$), more 20-second epochs scored as wake (Mantel-Haenszel $\chi^2 = 62.1$, $df = 1$, $p < 0.001$), and more 4-second epochs scored as wake (31% versus 16%, Mantel-Haenszel $\chi^2 = 271.4$, $df = 1$, $p < 0.001$). Summaries of total sleep by scoring method are given in Table 3. In the Kaplan-Meier analysis of the sleep continuity models (time till continuous sleep), the 8 continuous epoch(32-second) and 15 continuous epoch (60-second) models established no group differ-

ences, while that for the 30 continuous epoch (2-minute) model did (Breslow $P^2 = 4.03$, $df = 1$, $p = 0.045$).

Compared to controls, patients had more artifacts (20.5% versus 12.4%) (Mantel-Haenszel $\chi^2 = 142.7$, $df = 1$, $p < 0.001$) across all epochs. The artifacts differences were also present after t_0 (Mantel-Haenszel $\chi^2 = 151.8$, $df = 1$, $p < 0.001$); however this difference was not present when the first matched pair were removed (Mantel-Haenszel $\chi^2 = 0.43$, $df = 1$, $p = 0.8$). With this pair removed, the artifacts frequencies were about the same (14.2% for patients versus 12.8% for controls). The first matched insomnia subject's record was an outlier for artifact frequency. The Wilcoxon test of the time distribution of epochs after t_0 containing artifacts with the first matched pair excluded supported no group differences ($Z = 0.54$).

Analysis of Cumulative Sleep

Linear regression analyses of sleep build-up starting from t_0 through the subsequent 420 epochs were all highly significant for time, group, and group-by time interactions (all $p < 0.001$) in all scoring methods. Of greater interest was whether group differences in cumulative sleep could be documented at very short times after t_0 . Figure 4 displays the mean number of accumulated epochs of sleep after good night time as scored in 4-second epochs in the two groups graphed from 60 seconds before through 60 seconds after t_0 . The tests of mixed effects models using the endpoints of 60 seconds, 32 seconds, and 20 seconds supported no differences in the group term. However, the likelihood ratios for the model tests adding stepwise the group-by-time interaction term were 2.85, 1.01,

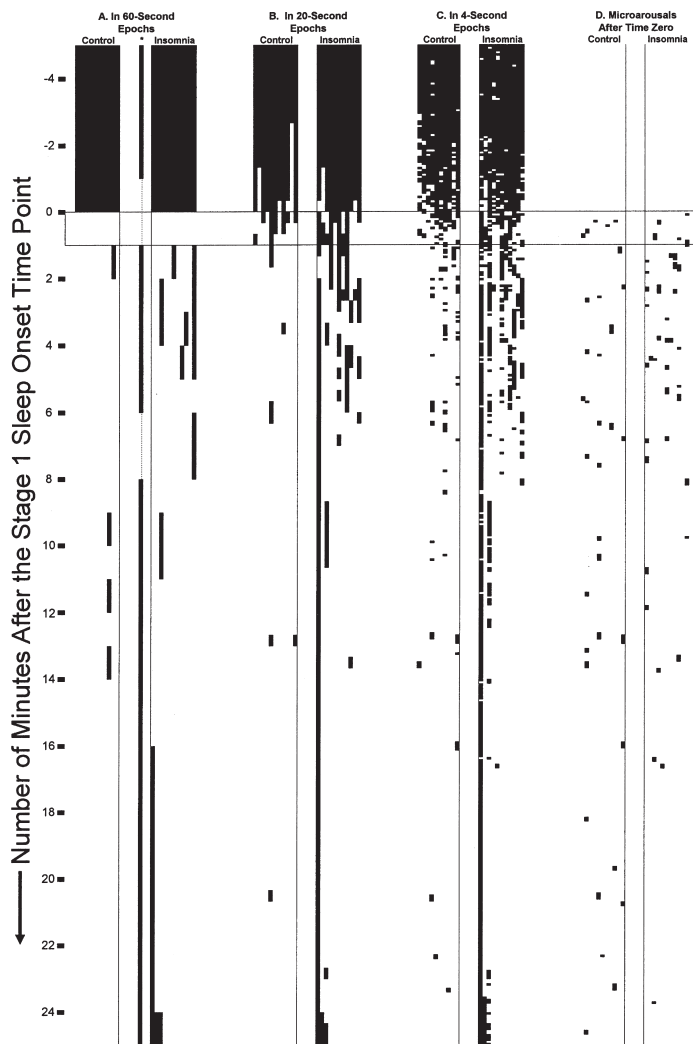


Figure 3—Scorings of the Sleep Onset Period Studied. In all displays, time proceeds from top to bottom starting from the beginning of the records analyzed. The horizontal boxed area represents the first 60-second epoch originally scored as stage 1 sleep in the archive. Time zero occurs at the beginning of this first 60-second epoch of stage 1 sleep. In each group block, the group (control vs. patient) is ordered from left to right in order of age. Darkened areas represent epochs scored as wake or aroused. A. Archival 60-second scoring of the record. The rescored record of the first insomnia patient is noted under “*” and is adjacent to the original record. B. Group-blind 20-second scoring of the record C. Group-blind 4-second scoring according to the newly developed classification rules (see Appendix). D. Group-blind ASDA arousal scoring after t_0 .

and 0.62 for the respective models ($df = 1$). Here only the model for the first 60-second epoch was significant ($p = 0.047$). The parallel models using cumulative 20-second epochs supported neither a group term nor a group-by-time interaction term (Likelihood ratio 1.91, $df = 1$, $p = 0.17$). Since the mixed modeling allowed for individual variability in intercepts and slopes, this approach was a more stringent test. The insomnia group had a statistically slower rate of entering sleep detectable with the 4-second scoring, but not with the 20-second scoring, at this 60-second time point.

Tests of ASDA Arousals

During the sleep onset period studied, there were no group differences found in the number or total time of ASDA arousals

Cumulative Epochs of Sleep

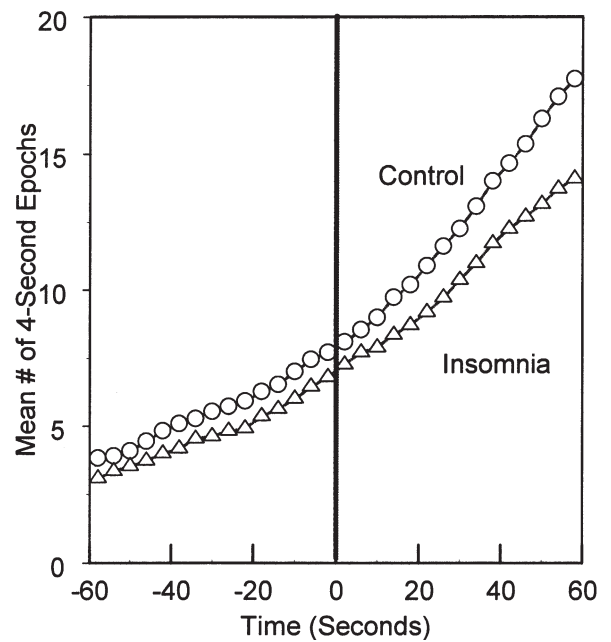


Figure 4—Average cumulative epochs of sleep scored in 4-second epochs as time-referenced to t_0 .

or long arousals. However, the Wilcoxon test of the distribution of ASDA arousals showed that patients had their ASDA arousal events earlier ($W = 1840$, $n = 39$, $p = 0.03$). With Kaplan Meier analysis, no group differences in the time until the first ASDA arousal after t_0 were found. However, group differences were present for the interval between the first and second ASDA arousals ($\chi^2 = 3.9$, $df = 1$, $p = 0.047$), but not for the interval between the second and third ASDA arousals.

Exploratory Correlations with Self-Report Data

The time interval data was correlated against all the post-sleep self-report variables listed in Table 2. The only comparisons found to be significant were those between the t_0 to 20-second sleep latency interval and the items asking about morning confusion, anxiety, and alertness ($r = -0.69, -0.75, -0.62$, respectively). Upon graphical inspection, these correlations were noted to be caused by 2 outlier values, and thus were not confidently supportable for the number of patients available.

DISCUSSION

Our primary aim was to describe sleep onset process abnormalities in primary insomnia patients in ways distinct from conventional sleep latency analysis. In conventional sleep latency analysis, the endpoint is a time point somewhere in stable stage 2 sleep, well past the events of the sleep onset process. In large part, this study provided comparisons that addressed disturbances before the sleep latency point, as here defined by the 20-second scoring rule definition. Since no one metric would suffice to summarize a process or its abnormalities, we employed a variety of methods of analyses to bring forth findings of sleep onset process abnormalities in the insomnia patients. The findings included group differences in visually assessed epoch patterns seen in Figure 3, differences in the number of 4-second epochs after t_0 ,

Table 2—Self-Reports Concerning the Night of Study

Sleep Parameters	Controls mean (SD)	Patients mean (SD)	<i>t</i> (df)¹	<i>p</i>
How long did it take you to fall asleep last night? (minutes)	13.1 (10.0)	38.1 (41.3)	2.00 (10)	0.073
How long did it take you to fall asleep usually? (minutes)	8.5 (8.1)	50.0 (43.7)	2.95 (10)	0.015
How long did you sleep last night? (hours)	6.9 (0.6)	5.8 (1.5)	-3.73 (10)	0.005
How long did you sleep usually? (hours)	7.1 (0.4)	5.8 (1.8)	-3.05 (10)	0.014
How many times did you awaken last night? (times)	1.9 (1.2)	2.8 (1.4)	1.70 (7)	0.133
How many times did you awaken usually? (times)	1.2 (0.9)	2.7 (1.3)	3.40 (10)	0.007
How long were you awake during the entire night last night? (minutes)	29.9 (37.0)	75.9 (66.8)	1.43 (11)	0.190
How long were you awake during the entire night usually? (minutes)	12.3 (18.9)	63.3 (46.6)	3.43 (7)	0.011
Visual-Analogue Scale Responses (0...100 mm Scale)				
Last night I felt very sleepynot at all sleepy.	30.6 (26.4)	35.6 (24.1)	0.42 (10)	0.687
Last night I had great difficulty falling asleepno difficulty.	70.9 (26.6)	61.2 (22.5)	-0.73 (10)	0.481
Last night I slept not at all soundlyvery soundly.	71.5 (22.0)	51.5 (24.2)	-1.64 (10)	0.132
Last night I slept less than I needas much as I need.	79.9 (22.0)	46.0 (28.6)	-2.41 (10)	0.037
This morning I had great difficulty awakeningno difficulty.	92.7 (4.7)	77.2 (21.6)	-2.14 (9)	0.061
This morning I feel poorly restedwell rested.	81.8 (23.7)	51.0 (25.3)	-2.27 (10)	0.047
This morning I feel very confusednot at all confused.	94.2 (5.4)	77.7 (20.5)	-2.62 (10)	0.026
This morning I feel very depressednot at all depressed.	94.8 (4.2)	82.0 (14.3)	-2.84 (10)	0.018
This morning I feel very anxiousnot at all anxious.	87.8 (22.0)	75.3 (22.2)	-1.18 (10)	0.266
This morning I feel not at all alertvery alert.	83.7 (22.7)	61.2 (27.9)	-1.70 (10)	0.119

¹Paired *t*-tests on available matched control-patient pairs.

differences in the timing of ASDA arousals after t_0 , differences in times until 2 minutes' continuous sleep in 4-second epochs after t_0 , differences in 4-second sleep latencies, differences in intervals between t_0 and the 4-second sleep latency point, and differences in the rate of accumulating sleep after t_0 within just 60 seconds after t_0 . For each endpoint used, some doubt might have remained that one had not selected the correct endpoint for a process analysis. Considered collectively, however, the coherence of the positive findings point to a disturbed sleep onset process in primary insomnia, in a sample not yet large enough to show robust sleep latency differences between groups. Sampling had not been based upon sleep latency characteristics.

The present exploratory study contributes to an existing literature^{6, 10, 15, 16, 19-23} on sleep onset process abnormalities in primary insomnia, by using methods of analysis not previously explored. Overall, the results confirmed the proposal of Saper¹³ that there may be faulty sleep/wake switching mechanisms in insomnia. The sleep onset process is a clinically relevant time domain where faulty sleep switch mechanisms may be studied. Saper's perspective that the difficulties may be in a moment-to-moment time domain (see Figures 3 and 4) appears to be correct, at least when

viewed from an approximate midpoint (t_0) of the sleep onset process. The difficulty at present is that no one approach to studying the sleep onset process seems to be obviously the scientifically best one, even though some, such as analysis of ASDA arousals, are less time consuming.

These results point to the value of analyzing the sleep onset process from a mid-process time referencing perspective. The sleep onset process orientation of t_0 in this study was to a process midpoint. The high correlations between t_0 latency and the later 20-second sleep latencies generally confirmed this midpoint perspective. There could be drawbacks to latency definitions insofar as the starting time point ("good night time") is determined by technicians rather than the patients themselves, and this starting

Table 3—Average minutes asleep as estimated by different scoring rules (all subjects included)

Scoring Rule	60-sec	20-sec	4-sec
Control	27.6	27.8	27.6
Insomnia	25.2	23.2	23.3

point determination has unknown reliability. The midpoint method is more amenable to reliability testing. All the data for setting the midpoint is on the PSG record itself, rather than depending on external information.

The sleep latency definition that we applied to the 4-second epoch data had the similarity of a 10-minute perspective to those sleep latency definitions applied to the 60-second and 20-second data, but was dissimilar in that it had no linkage to stable stage 2 sleep. Some time in stage 2 sleep is needed for a sleep latency point to be reasonably defined. When t_0 and the 4-second definitions of sleep latency were co-plotted (see Figure 2), they appeared similar. The new 4-second definition of “sleep latency” was not comparable to usual definitions of sleep latency. In retrospect, the 4-second sleep latency point might itself have served as the t_0 for this analysis. The non-equivalency of sleep latency definitions is an embedded problem in sleep onset process research,¹⁷ especially when scoring is done in different epoch lengths.

The findings that insomnia patients can have only mild or no changes in all-night conventional sleep architecture^{3, 4, 7} can make it seem that these patients only have misperceptions about sleep and sleep latency. In this study, the patients reported only statistical trend differences in sleep latencies compared to the controls for the night studied. The accuracy of the patients’ self-reports was not grossly abnormal. By virtue of their group membership, the patients’ self-reports had some general correspondence to sleep onset process abnormalities that we could document compared with the matched control subjects. Three correlations between the sleep onset interval in the patients and their self-reported confusion, anxiety, and alertness levels the following morning seemed conceptually promising, but cannot be considered more than possible suggestions for future testing. At the individual patient level, self-report variables did not have fixed and precise relationships to physiologically derived variables.

Three issues are worth noting when considering new scoring rules for insomnia research. First, scorers use visual pattern recognition to score epochs. They filter out the pattern of artifacts and/or other phenomena when scoring the record. A recent publication suggests that even for staging of 30-second epochs, specific sleep stages have subtypes that the scorers gloss over while scoring.³⁶ Patients may have more movement arousals during sleep,^{27, 28} which might be considered as artifacts. Hence, one can be skeptical that current scoring methods are optimal for studying insomnia. Is the “misperception of sleep” more in the scoring procedures or in the patients? Second, conventional epoch lengths, while convenient, are not necessarily the best ones for characterizing the sleep of insomnia patients. Third, if the “independent” presence of artifacts is related to group or subtype differences, then artifact scoring may provide useful information. In the present study, the one insomnia subject’s increased artifacts may have been due to purely technical factors, or may represent a physiological subtype. Analyzing artifact patterns will help avoid potential biases.

The present study had a specialized case-control design which involved individual matching in a small sample. Being a case control study, sampling was based on the outcome the subjects had already experienced, and not with random sampling. Hence, the methods and findings from the study can only be taken as exploratory.

This study may be the only report where scorings of different epoch lengths have been compared directly. Further research studies of scoring sleep and sleep onset across different epoch sizes may be helpful in examining factors affecting scoring reliability.

The rules used in this study for scoring 4-second epochs will need further replication, testing, and modification. Using them may help to refine theories about the sleep onset process. Wider application of the Hori scoring rules may also be important. Other small-epoch scoring rules are conceivable and potentially useful. Yet since scoring small epochs requires a larger amount of scoring time, automated scoring will likely be needed if small-epoch scoring is to have any practical clinical application.

Small-epoch scoring is unlikely to be of much practical clinical use unless it can contribute to a differential diagnosis of insomnia subtypes. If there are subtypes of sleep onset process abnormalities, as suggested by the results of Staner et al,²¹ then small-epoch scoring may have clinical utility. To accomplish such a goal of differential diagnosis using polysomnography, clinical investigators will need to posit and test potential physiological subtypes, and not rely merely on the current nonspecific, polythetic DSM-IV³⁷ definition of primary insomnia or the ICSD-2³⁸ definitions of chronic insomnias. Using the current definitions may only produce study samples that are confounded mixtures of physiological subtypes. The present study’s sample was selected on the basis that the DSM-IV definition served as the outcome variable that determined the sampling frame for the case-control design. Not surprisingly, there were some insomnia subjects with comparatively normal sleep latencies in the sample. Future studies should focus on potential physiologically defined subtypes of the sleep onset process to set up case-control studies, rather than focusing primarily on syndromal outcomes that are mainly defined by self-reports. By and large, patients’ self-reports describe well the symptoms, but not the physiological causes, of their diseases.

ACKNOWLEDGMENTS

Research funding from NIMH grant MH24652, and National Institute of Aging grants AG00972 and AG20677, and NIH/NCRR/GCRC Grant M01-RR000056 supported this work. We thank Robert Ogilvie, Ph.D., Cindy Ehlers, Ph.D., Raymond Vasiko, Ph.D., and Patricia Houck, M.A., for technical assistance.

REFERENCES

1. Steriade M, McCarley RW. Brain Control of Wakefulness and Sleep. 2nd ed. ed. New York: Springer; 2005.
2. Rechtschaffen A, Kales A. A Manual of Standardized Terminology, Techniques, and Scoring Systems of Sleep Stages in Human Subjects. Los Angeles, CA: UCLA Brain Information Service / Brain Research Institute; 1968.
3. Rosa RR, Bonnet MH. Reported chronic insomnia is independent of poor sleep as measured by electroencephalography. *Psychosom Med* 2000;62:474-82.
4. Edinger JD, Fins AI, Glenn DM, et al. Insomnia and the eye of the beholder: are there clinical markers of objective sleep disturbances among adults with and without insomnia complaints? *J Consult Clin Psychol* 2000;68:586-93.
5. Fichten CS, Creti L, Amsel R, Brender W, Weinstein N, Libman E. Poor sleepers who do not complain of insomnia: myths and realities about psychological and lifestyle characteristics of older good and poor sleepers. *J Behav Med* 1995;18:189-223.
6. Merica H, Blois R, Gaillard JM. Spectral characteristics of sleep EEG in chronic insomnia. *Eur J Neurosci* 1998;10:1826-34.
7. Perlis ML, Giles DE, Mendelson WB, Bootzin RR, Wyatt JK. Psychophysiological insomnia: the behavioural model and a neurocognitive perspective. *J Sleep Res* 1997;6:179-88.

8. Perlis ML, Merica H, Smith MT, Giles DE. Beta EEG activity and insomnia. *Sleep Med Rev* 2001;5:365-76.
9. Perlis ML, Smith MT, Andrews PJ, Orff H, Giles DE. Beta/Gamma EEG activity in patients with primary and secondary insomnia and good sleeper controls. *Sleep* 2001;24:110-7.
10. Merica H, Gaillard JM. The EEG of the sleep onset period in insomnia: a discriminant analysis. *Physiol Behav* 1992;52:199-204.
11. Rodenbeck A, Rütther E, Cohrs S, Hajak G. Quantifizierte Arousal-Analyse bei Patienten mit einer psychophysiologischen Insomnie. *Somnologie* 2000;4:55-60.
12. Borkovec TD, Lane TW, VanOot PH. Phenomenology of sleep among insomniacs and good sleepers: wakefulness experience when cortically asleep. *J Abnorm Psychol* 1981;90:607-9.
13. Saper CB, Chou TC, Scammell TE. The sleep switch: hypothalamic control of sleep and wakefulness. *Trends Neurosci* 2001;24:726-31.
14. Levitt H, Wood A, Moul DE, et al. A pilot study of subjective daytime alertness and mood in primary insomnia participants using ecological momentary assessment. *Behav Sleep Med* 2004;2:113-31.
15. Freedman RR. EEG power spectra in sleep-onset insomnia. *Electroencephalogr Clin Neurophysiol* 1986;63:408-13.
16. Bonato R. Electroencephalographic Correlates of Sleep Onset in Chronic Psychophysiological Insomniacs and Normal Sleepers [Ph. D. Thesis]. Ottawa, Canada: Carleton University; 1997.
17. Ogilvie RD. The process of falling asleep. *Sleep Med Rev* 2001;5:247-70.
18. Bonnet MH, Webb WB. The return to sleep. *Biol Psychol* 1979;8:225-33.
19. Jacobs G, Benson H, Friedman R. Home-based central nervous system assessment of a multifactor behavioral intervention for chronic sleep-onset insomnia. *Behav Ther* 1993;24:159-74.
20. Lamarche CH, Ogilvie RD. Electrophysiological changes during the sleep onset period of psychophysiological insomniacs, psychiatric insomniacs, and normal sleepers. *Sleep* 1997;20:724-33.
21. Staner L, Cornette F, Maurice D, et al. Sleep microstructure around sleep onset differentiates major depressive insomnia from primary insomnia. *J Sleep Res* 2003;12:319-30.
22. Perlis ML, Kehr EL, Smith MT, Andrews PJ, Orff H, Giles DE. Temporal and stagewise distribution of high frequency EEG activity in patients with primary and secondary insomnia and in good sleeper controls. *J Sleep Res* 2001;10:93-104.
23. Williams BR. Electrophysiological analysis of the sleep onset period: A comparison between subjects with long term insomnia complaints associated with mild traumatic brain injury and matched controls [Masters Thesis]. St. Catharines, Ontario: Brock University; 1999.
24. Doerfling P, Ogilvie RD, Murphy T, Lamarche C. Applying the Hori sleep scoring system to the examination of the sleep onset process in insomniac and normal sleepers. *Sleep Res* 1996;25:123.
25. Hori T, Hayashi M, Morikawa T. Topographical EEG changes and the hypnagogic experience. In: Ogilvie RD, Harsh JR, editors. *Sleep Onset: Normal and Abnormal Processes*. Washington, DC: American Psychological Association; 1994. p. 237-53.
26. Chesson A, Jr., Hartse K, Anderson WM, et al. Practice parameters for the evaluation of chronic insomnia. An American Academy of Sleep Medicine report. Standards of Practice Committee of the American Academy of Sleep Medicine. *Sleep* 2000;23:237-41.
27. Polo-Kantola P, Erkkola R, Irjala K, Pullinen S, Virtanen I, Polo O. Effect of short-term transdermal estrogen replacement therapy on sleep: a randomized, double-blind crossover trial in postmenopausal women. *Fertil Steril* 1999;71:873-80.
28. Bonnet MH, Arand DL. Impact of activity and arousal upon spectral EEG parameters. *Physiol Behav* 2001;74:291-8.
29. American Psychiatric Association, American Psychiatric Association. Task Force on DSM-IV. *Diagnostic and Statistical Manual of Mental Disorders: DSM-IV*. 4th ed. Washington, DC: American Psychiatric Association; 1994.
30. First M, Spitzer RL, Gibbon M, Williams JBW. Structured Clinical Interview for DSM-IV Axis I Disorders - Patient Edition (SCID-I/P). New York: New York State Psychiatric Institute; 1995.
31. American Sleep Disorders Association. *The International Classification of Sleep Disorders: Diagnostic and Coding Manual*. Rochester, MN: American Sleep Disorders Association; 1990.
32. Sleep Disorders Atlas Task Force of the American Sleep Disorders Association. EEG arousals: scoring rules and examples. *Sleep* 1992;15:174-84.
33. SPSS for Windows, Version 11.5. In: Standard Release ed. Chicago, IL: SPSS Inc.; 2002.
34. S-PLUS 7 for Windows. In: Seattle, WA: Insightful Corp.; 2005.
35. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res* 1989;28:193-213.
36. Muller B, Gabelein WD, Schulz H. A taxonomic analysis of sleep stages. *Sleep* 2006;29:967-74.
37. American Psychiatric Association., American Psychiatric Association. Task Force on DSM-IV. *Diagnostic and statistical manual of mental disorders: DSM-IV-TR*. 4th ed. Washington, DC: American Psychiatric Association; 2000.
38. American Academy of Sleep Medicine. *The International Classification of Sleep Disorders: Diagnostic & Coding Manual*. 2nd ed. Westchester, IL: American Academy of Sleep Medicine; 2005.

APPENDIX

Classification Routine for 4-Second Epochs

This appendix describes a system of classification that uses two primary facets of the signal and classifies 4-second epochs of the central (C₄/C₃) EEG leads.

First, a 4-second epoch will be classified as **NREM**, **REM**, or **AWAKE**. Such classification requires a minimum of 50% of the epoch (2 seconds or greater) meeting amplitude and frequency requirements consistent with those of the Rechtschaffen and Kales criteria. Second, the classification determines whether it is **clean** or contains an **artifact**. It will be classified as containing an artifact if any of the following occur in its EEG, EOG or EMG channels, in accordance with the guidelines given below:

- EMG elevation/twitches
- Eye blink/sharp eye movement
- Slow eye movements
- EKG spike artifacts
- Sweat artifact
- Electrode pop in EEG/EOG channels
- 60 CPS in EEG, EOG, or EMG channel

An epoch's classification may require consideration of the surrounding epochs and transitions from AWAKE to NREM and NREM to REM.

Guidelines Regarding Artifacts

A. **EMG Artifacts:** To classify an EMG artifact requires that the scorer establish an EMG amplitude baseline from the previous epoch. Using that baseline reference is needed in relation to the following guidelines:

1. If the EMG amplitude of the epoch being scored increases by a factor of 4 for 0.5–2 seconds, the epoch is classified as containing an artifact.
2. Baseline EMG should be established by the last stable amplitude of 2 seconds from the previous epoch. Normally this will be the last 2 seconds of the previous

epoch. However, if this segment has a deviation from baseline of at least 0.5 seconds (i.e., muscle twitch), and the amplitude increases by a factor of 4, the immediately preceding 2-second segment should be considered to establish a baseline. One should review previous 2-second segments until a stable baseline is established. The first prior-qualifying epoch will then be the baseline EMG for comparison with the EMG of the epoch being scored.

3. An EMG spike that can be identified at the same time point in either the EEG or EOG channels should be classified as an artifact. However, if the spike is ≤ 0.5 seconds and cannot be identified in adjacent channels, the epoch is classified as clean.

B. Eye Movement Artifacts:

1. Eye blinks during wake and phasic eye movements during sleep often show a clear influence on the EEG. Visually identifying this influence and assessing the effect on the EEG can be tedious. Because of this, we have elected to classify all sharp eye movements as artifact when 1) the frequency is ≥ 1 Hz, and 2) the presence of an apex is clear. Therefore, the eyes' influence on the EEG is not required to classify the epoch as artifact. If the eye movement is sharp according to above it should be classified as artifact regardless of any notable influence on the EEG. The degree of influence on the EEG is often correlated with the eyes' deflection amplitude, but again this influence is not necessary for a classification of artifact. When the consideration point falls on the border/intersection of two epochs, the classification should be the same for both epochs.
2. Transient periods between wake–sleep states characteristically show slow rolling eye movements. One should classify slow rolling eyes as clean if; 1) the frequency is less than 1 Hz, and 2) the presence of a clear apex is absent.

C. Pulse/EKG Artifact: Pulse/EKG artifact can infiltrate EEG and EOG signals. When pulse artifact is present in the EEG the epoch should be classified as artifact. Generally this will persist throughout all epochs until the reference point or electrode is changed.

D. Sweat Artifact: Sweat artifact can affect EEG and EOG signal integrity and baseline. This is usually caused by reference electrodes and therefore will be seen in all EEGs and EOGs as a “floating” signal. When this artifact is present, the epoch should be classified as containing an artifact.

E. Electrode Pops/60 Hz artifacts: Reference and ground electrodes can corrupt the integrity of EEG and EOG signal. This effect can be subtle or sudden, so special attention to the point of origin is essential.