

Sleep Classification According to AASM and Rechtschaffen & Kales: Effects on Sleep Scoring Parameters

Doris Moser, PhD, MSc¹; Peter Anderer, PhD²; Georg Gruber, MSc³; Silvia Parapatics, PhD, MSc³; Erna Loretz, MSc³; Marion Boeck¹; Gerhard Kloesch¹; Esther Heller⁴; Andrea Schmidt⁴; Heidi Danker-Hopfe, PhD⁴; Bernd Saletu, MD²; Josef Zeithofer, MD, PhD¹; Georg Dorffner, PhD^{3,5}

¹Department of Neurology, Medical University of Vienna, Vienna, Austria; ²Department of Psychiatry and Psychotherapy, Medical University of Vienna, Vienna, Austria; ³The Siesta Group Schlafanalyse GmbH, Vienna, Vienna, Austria; ⁴Department of Psychiatry and Psychotherapy, Charité – Universitätsmedizin Berlin, Campus Benjamin Franklin (CBF), Berlin, Germany; ⁵Institute of Medical Cybernetics and Artificial Intelligence, Center for Brain Research, Medical University of Vienna, Vienna, Austria

Study Objective: To investigate differences between visual sleep scoring according to the classification developed by Rechtschaffen and Kales (R&K, 1968) and scoring based on the new guidelines of the American Academy of Sleep Medicine (AASM, 2007).

Design: All-night polysomnographic recordings were scored visually according to the R&K and AASM rules by experienced sleep scorers. Descriptive data analysis was used to compare the resulting sleep parameters.

Participants: Healthy subjects and patients (38 females and 34 males) aged between 21 and 86 years.

Interventions: N/A

Measurement and Results: While sleep latency and REM latency, total sleep time, and sleep efficiency were not affected by the classification standard, the time (in minutes and in percent of total sleep time) spent in sleep stage 1 (S1/N1), stage 2 (S2/N2) and slow wave sleep (S3+S4/N3) differed significantly between the R&K and the AASM classification. While light and deep sleep increased (S1 vs. N1 [+10.6 min, +2.8%]; $P < 0.01$; S3+S4 vs. N3 [+9.1 min (+2.4%)]; $P < 0.01$), stage

2 sleep decreased significantly according to AASM rules (S2 vs. N2 [-20.5 min, (-4.9%)]; $P < 0.01$). Moreover, wake after sleep onset was significantly prolonged by approximately 4 minutes ($P < 0.01$) according to the AASM standard. Interestingly, the effects on stage REM were age-dependent (intercept at 20 years: -7.5 min; slope: 1.6 min for 10-year age increase). No effects of sex and diagnosis were observed.

Conclusion: The study shows significant and age-dependent differences between sleep parameters derived from conventional visual sleep scorings on the basis of R&K rules and those based on the new AASM rules. Thus, new normative data have to be established for the AASM standard.

Keywords: Sleep scoring parameters, AASM standard, Rechtschaffen & Kales, sleep architecture, age.

Citation: Moser D; Anderer P; Gruber G; Parapatics S; Loretz E; Boeck M; Kloesch G; Heller E; Schmidt A; Danker-Hopfe H; Saletu B; Zeithofer J; Dorffner G. Sleep classification according to AASM and rechtschaffen & kales: effects on sleep scoring parameters. *SLEEP* 2009;32(2):139-149.

FOR APPROXIMATELY 40 YEARS THE ONLY WIDELY ACCEPTED STANDARD FOR DESCRIBING THE HUMAN SLEEP PROCESS WAS THE MANUAL OF SLEEP CLASSIFICATION by Rechtschaffen and Kales.¹ On the basis of these scoring rules, sleep recordings are divided into 7 discrete stages (wake, stage 1, stage 2, stage 3, stage 4, stage REM, and movement time). Even though in many cases this standard is useful, the rules of Rechtschaffen and Kales have also been criticized for leaving plenty of room for subjective interpretation, which leads to a great variability in the visual evaluation of sleep stages.^{2,3} Last but not least, the standard rules were developed for young healthy adults^{4,5} and do not necessarily directly apply to elderly subjects and patients.

The American Academy of Sleep Medicine (AASM)⁶ modified the standard guidelines for sleep classification by Rechtschaffen and Kales and developed a new guideline for terminology, recording method, and scoring rules for sleep-related phenomena. The manual is the result of a review of literature, analysis and consensus which addresses 7 topics: digital analy-

sis and reporting parameters, visual scoring, arousal, cardiac and respiratory events, movements and pediatric scoring. One of the major changes is a change in terminology: in the AASM classification, sleep stages S1 to S4 are referred to as N1, N2, and N3, with N3 reflecting slow wave sleep (SWS, R&K stages S3 + S4); stage REM is referred to as stage R. According to the AASM manual, a minimum of 3 EEG derivations, sampling activity from the frontal, central, and occipital regions, has to be recorded. The recommended derivations are F4-M1, C4-M1, and O2-M1 (right-sided active electrodes and a reference over the left mastoid, rather than the ear).⁷ The new manual also deals with the definition of the sleep-wake transition, sleep spindles, K-complexes, slow wave sleep, and REM sleep, as well as arousals and major body movements. In summary, the major changes of the new manual comprise EEG derivations, the merging of stages 3 and 4 into N3, the abolition of stage “movement time,” the simplification of many context rules as well as the recommendation of sampling rates and filter settings for polysomnographic (PSG) reporting and for user interfaces of computer-assisted sleep analysis.⁶

To date there are no studies evaluating the effects of the new standard on sleep scoring data. The aim of the present investigation was to describe in detail differences between visual sleep scoring according to the Rechtschaffen and Kales classification and scoring based on the new AASM guidelines in normal subjects of different age groups and sleep-disturbed patients.

Submitted for publication April, 2008

Submitted in final revised form October, 2008

Accepted for publication November, 2008

Address correspondence to: Univ. Prof. Dr. Josef Zeithofer, Department of Neurology, Medical University of Vienna, Waehringerguertel 18-20, 1090 Vienna, Austria; Tel: 0043-1-40400-3100; Fax: 0043-1-40400-3141; E-mail: josef.zeithofer@meduniwien.ac.at

EXPERIMENTAL DESIGN AND METHODOLOGY

This study is based on data of the database recorded in the SIESTA project (“A New Standard for Integrating Polygraphic Sleep Recordings into a Comprehensive Model of Human Sleep and its Validation in Sleep Disorders”).⁸

Subjects

Nocturnal sleep recordings of 72 subjects (38 females and 34 males) aged between 21 and 86 years (mean age 58 ± 19 years) were analyzed. The sample consisted of 56 healthy subjects (mean age 58 ± 20 years), 5 patients with general anxiety disorder (GAD; mean age 41 ± 12 years), 5 patients with periodic limb movement disorders (PLMD; mean age 55 ± 16 years) and 6 patients with Parkinson disease (mean age 66 ± 1 years).

Polysomnography

Each subject underwent polysomnographic all-night recordings (PSG) in the sleep laboratory for 2 consecutive nights (day 7 and 8 of a 14-day study period). For this study, only the PSGs of the second night were analyzed. All recordings started at the subjects’ usual bedtime and ended at their usual time of getting up in the morning. PSG recordings included 19 EEG channels, 2 electrooculogram (EOG) channels, submental electromyogram (EMG) and EMG recorded from electrodes placed at the musculus anterior tibialis of the left and right leg, electrocardiogram (ECG) and respiratory signals (airflow, movements of the chest wall and abdomen, O₂ saturation of arterial blood). In the recordings used in the present study, EEG, EOG, and EMG signals were digitized with 200 Hz or 256 Hz. The high-pass filters for EEG and EOG recordings were between 0.16 Hz and 0.5 Hz; for EMG, 1.6 Hz; and for ECG, 16 Hz. Respiratory signals were sampled with 16 Hz, and O₂ saturation with 1 Hz.^{8,9}

Data Analysis

Visual analyses of 72 PSGs based on the standard rules by Rechtschaffen and Kales (1968) and the new AASM rules (2007) were compared.

During the SIESTA project, PSGs were visually scored by 2 independent scorers out of a pool of 30 sleep experts from 8 European sleep labs. All 7 (sleep) stages (wake, sleep stage 1 [S1], sleep stage 2 [S2], sleep stage 3 [S3], sleep stage 4 [S4], sleep stage REM [REM] and movement time [MT]) were scored according to the standard guidelines developed by Rechtschaffen and Kales (1968). Thus, the scoring was based on 2 EOG, one (sub)mental EMG and the central EEG channels. If the scorings were not concordant, a third scorer took the final “consensus” decision. In the present analysis, the consensus scorings were used.

In August 2007, the same 72 PSGs were visually scored by 7 sleep experts from 3 European sleep labs. All PSGs were visually scored by 2 independent scorers, 12 out of the 72 PSGs were classified by 6 scorers. (Sleep) stages [wake (W), sleep stage 1 (N1), sleep stage 2 (N2), sleep stage 3 (N3) and sleep stage REM (R)] were scored according to the new AASM scoring rules. AASM scoring was preceded by a 2-day training

symposium with detailed discussions and 4 test scorings (not included in the present study) to ensure that all scorers correctly interpreted the new AASM standard. Note that stage N3 represents slow wave sleep and corresponds to the Rechtschaffen and Kales stages S3 and S4. In the present analysis, the first scorings were used.

To estimate the influence of interrater variability on the results of the present study, the range for the differences between AASM and R&K scorings for all possible pair-wise comparisons between the available independent expert scorings are given (AASM 1st scoring versus R&K 1st scoring, AASM 1st scoring versus R&K 2nd scoring, AASM 2nd scoring versus R&K 1st scoring, AASM 2nd scoring versus R&K 2nd scoring).

Statistical Analysis

All values are reported as means \pm standard deviations. SPSS version 14.0 (SPSS, Inc., Chicago, IL, USA) was used for analysis. Sleep scoring data derived from R&K and AASM standards were tested for significant differences by means of a paired-samples *t*-test for normally distributed data or in case of a violation of the normal distribution by a Wilcoxon matched-pairs signed-ranks test. The Kolmogorov-Smirnov one-sample test was applied to test for a violation of normal distribution. $P < 0.05$ was considered significant. To test whether the changes in the sleep data derived from the R&K and the AASM standard are affected by sex and age, ANOVAS with “classification standard” as repeated measurement factor and sex and age as independent factors was computed for the healthy subjects. Moreover, to evaluate a possible interaction between “classification standard” and diagnosis, ANOVAS with classification standard as repeated measurement factor and diagnosis as independent factor were computed for patients and age- and sex-matched healthy subjects. In addition, the change values of the sleep data derived from AASM and R&K standards were computed (“AASM – R&K”) and these change values were tested for differences between females and males, young and older subjects as well as patients and matched controls by independent-samples *t*-tests or Mann-Whitney U-tests for normally and non-normally distributed data, respectively. Moreover, regression analyses with the change values “AASM – R&K” as dependent variable and age as independent variable were performed. Finally, multiple regression analysis with age and age-square as independent variables was performed to reveal potential nonlinearities.

RESULTS

Total Group

Table 1 provides summary statistics of the demographic data, including age, body mass index (BMI), and sleep quality (Pittsburgh Sleep Quality Index, PSQI)¹⁰ stratified by sex.

Table 2 summarizes the results for all sleep parameters, as recommended by the AASM Manual.⁶ The sleep parameters derived from R&K and AASM scorings were compared by means of Student’s *t*-tests for normally distributed variables and by Wilcoxon tests for not normally distributed variables. Throughout the paper, P-values for sleep latency (SL), stage R

Table 1—Demographic Data (Mean ± Standard Deviation)

	N	Age (years)	BMI (kg ² /m)	PSQI (score)
Total group	72	57.68 ± 18.72	25.33 ± 3.70	4.93 ± 3.25
Male	34	59.76 ± 17.86	26.20 ± 2.35	5.36 ± 3.16
Female	38	55.82 ± 19.50	24.58 ± 4.45	4.54 ± 3.32
Healthy subjects	56	58.48 ± 19.76	25.58 ± 3.56	3.63 ± 1.25
Male	25	59.72 ± 20.00	26.52 ± 2.31	3.88 ± 1.08
Female	31	57.48 ± 19.84	24.86 ± 4.18	3.43 ± 1.36
Healthy subjects < 60 years	25	39.16 ± 10.97	24.15 ± 3.97	3.37 ± 1.24
Male	11	40.64 ± 1.78	25.82 ± 2.69	3.73 ± 1.35
Female	14	38.00 ± 10.59	22.86 ± 4.40	3.08 ± 1.12
Healthy subjects ≥ 60 years	31	74.06 ± 7.58	26.68 ± 2.82	3.83 ± 1.23
Male	14	74.71 ± 9.03	27.06 ± 1.90	4.00 ± 0.82
Female	17	73.53 ± 6.39	26.39 ± 3.39	3.71 ± 1.49
Patient group	16	54.87 ± 14.70	24.51 ± 4.12	9.31 ± 4.06
Male	9	59.89 ± 10.78	25.38 ± 2.38	9.33 ± 3.50
Female	7	48.43 ± 17.30	23.39 ± 5.68	9.29 ± 4.99
Age- and sex-matched healthy subjects	16	55.06 ± 14.37	25.64 ± 3.56	3.27 ± 1.16
Male	9	59.89 ± 10.47	26.94 ± 2.20	3.75 ± 0.71
Female	7	48.89 ± 17.04	24.35 ± 4.33	2.71 ± 1.38

Table 2—Sleep Scoring Parameters (Mean ± Standard Deviation) Derived from Sleep Classification According to R&K and AASM for the Total Group (n = 72)

	R&K	AASM	Difference AASM-R&K	P-value
Lights out clock time (h:min)	23:03 ± 0:26	23:03 ± 0:26	-	-
Lights on clock time (h:min)	07:06 ± 0:32	07:06 ± 0:32	-	-
Total sleep time (TST; in min)	397.1 ± 56.7	394.5 ± 56.6	-2.5	0.107
Total recording time				
("lights out" to "lights on" in min)	481.3 ± 33.4	481.3 ± 33.4	-	-
Sleep latency (SL; in min)*	16.0 ± 13.1	15.7 ± 14.3	-0.4	0.775
Sleep REM/R latency				
(sleep onset to first epoch of Stage R in min)	84.2 ± 40.7	88.4 ± 40.9	4.2	0.143
Wake after sleep onset				
(WASO; Stage W after sleep onset, in min)	65.9 ± 45.8	69.9 ± 45.8	4.1	0.008
Percent sleep efficiency				
(TST/total recording time)x100 in %	82.7 ± 10.3	82.2 ± 10.5	-0.5	0.110
Time in each stage (min)				
Stage S1/N1	48.8 ± 23.5	59.4 ± 29.2	10.6	0.000
Stage S2/N2	215.9 ± 46.8	195.4 ± 48.4	-20.5	0.000
Stage S3+S4/N3	53.7 ± 32.5	62.7 ± 40.4	9.1	0.006
Stage REM/R	78.9 ± 22.3	77.1 ± 23.4	-1.8	0.107
Percent of TST in each stage				
(time in each stage/TST)x100 in %				
Stage S1/N1	12.4 ± 5.9	15.3 ± 7.7	2.8	0.000
Stage S2/N2	54.1 ± 7.6	49.2 ± 9.3	-4.9	0.000
Stage S3+S4/N3	13.7 ± 8.6	16.1 ± 10.8	2.4	0.005
Stage REM/R	19.7 ± 4.8	19.4 ± 5.1	-0.4	0.175

The P-values are based on Wilcoxon's matched pairs signed-ranks test for TST, SL, REM/R-latency, WASO and sleep efficiency, for all other variables on paired samples *t*-tests. Numbers in boldface indicate significant results.

*Definition: R&K: lights out to 3 consecutive epochs S1 or the first epoch of any deeper sleep.

AASM: lights out to first epoch of any sleep.

latency, and sleep efficiency are based on Wilcoxon matched-pairs signed-ranks tests; all other p-values are based on *t*-tests. Total sleep time (TST) differed by -2.5 min between AASM and R&K scorings (not significant (NS); range for all possible pairs of 1st and 2nd scorings was -1.7 to -8.2 min). Sleep latency (SL) differed by -0.4 min (NS; range: 0.5 to -1.8 min)

and REM/R latency by 4.2 min (NS; range: 5.8 to -0.3 min). Wake after sleep onset (WASO) differed significantly by 4.1 min (P = 0.008; range: 1.7 to 9.5 min). Sleep efficiency differed by -0.5% (NS; range: -0.0% to -1.7%). Time in stage S1/N1 increased significantly by 10.8 min (P = 0.000; range: 3.3 to 12.7 min), time in stage S2/N2 decreased significantly by -20.5

Table 3—Sleep Scoring Parameters (Mean ± Standard Deviation) Derived from Sleep Classification According to R&K and AASM for Young (Y: < 60 Years, n = 25) and Older (O: ≥ 60 Years, n = 31) Healthy Subjects

	Y: Healthy subjects: < 60 years				O: Healthy subjects: ≥ 60 years				Y – O	
	R&K	AASM	Diff.	P	R&K	AASM	Diff.	P	Diff.	P
			AASM-R&K				AASM-R&K		AASM-R&K	
Lights out clock time (hr:min)	23:15 ± 0:25	23:15 ± 0:25	-	-	23:01 ± 0:17	23:01 ± 0:17	-	-	-	-
Lights on clock time (hr:min)	07:16 ± 0:30	07:16 ± 0:30	-	-	07:04 ± 0:26	07:04 ± 0:26	-	-	-	-
Total sleep time (TST; in min)	422.3 ± 41.6	418.0 ± 40.1	-4.3	0.109	384.9 ± 56.4	382.9 ± 57.0	-2.0	0.241	-2.3	0.921
Total recording time (“lights out” to “lights on” in min)	478.1 ± 32.9	478.1 ± 32.9	-	-	481.0 ± 25.4	481.0 ± 25.4	-	-	-	-
Sleep latency (SL; in min)*	16.3 ± 13.7	14.2 ± 11.3	-2.1	0.435	16.5 ± 14.1	16.4 ± 15.0	-0.0	0.914	-2.1	0.493
Sleep REM/R latency (sleep onset to first epoch of Stage R in min)	82.7 ± 28.1	88.4 ± 34.4	5.7	0.085	77.4 ± 35.1	83.7 ± 44.6	6.3	0.705	-0.7	0.228
Wake after sleep onset (WASO; Stage W after sleep onset, in min)	39.8 ± 24.9	44.6 ± 28.0	4.8	0.074	74.8 ± 45.9	79.9 ± 46.2	5.1	0.036	-0.3	0.609
Percent sleep efficiency (TST/total recording time)x100 in %	88.0 ± 6.6	87.7 ± 6.5	-0.3	0.590	80.6 ± 10.5	79.9 ± 10.5	-0.7	0.088	0.4	0.525
Time in each stage (min)										
Stage S1/N1	42.8 ± 21.4	49.5 ± 23.7	6.7	0.042	54.4 ± 24.8	66.8 ± 32.0	12.4	0.001	-5.6	0.245
Stage S2/N2	224.1 ± 41.3	208.6 ± 43.3	-15.5	0.017	214.5 ± 46.8	186.3 ± 49.4	-28.2	0.000	12.7	0.178
Stage S3+S4/N3	63.2 ± 31.4	72.1 ± 38.2	8.9	0.091	47.0 ± 35.1	59.5 ± 44.1	12.5	0.027	-3.5	0.642
Stage REM/R	92.2 ± 16.7	87.8 ± 19.2	-4.4	0.032	69.1 ± 19.3	70.4 ± 22.0	1.3	0.357	-5.8	0.018
Percent of TST in each stage (time in each stage/TST)x100 in %										
Stage S1/N1	10.4 ± 5.5	12.1 ± 6.3	1.7	0.036	14.2 ± 6.2	17.6 ± 8.2	3.4	0.001	-1.7	0.187
Stage S2/N2	52.8 ± 7.0	49.7 ± 8.4	-3.1	0.027	55.6 ± 8.2	48.4 ± 10.5	-7.2	0.000	.0	0.072
Stage S3+S4/N3	14.9 ± 7.2	17.2 ± 8.9	2.3	0.078	12.4 ± 9.3	15.7 ± 11.7	3.3	0.026	-1.0	0.594
Stage REM/R	21.9 ± 3.5	21.0 ± 3.9	-0.9	0.061	17.8 ± 4.3	18.2 ± 5.2	0.4	0.239	-1.3	0.023

The P-values are based on Wilcoxon's matched pairs signed-ranks test (Y, O) or Mann-Whitney U-test (Y – O) for TST, SL, REM/R-latency, WASO and sleep efficiency, for all other variables on paired samples (Y, O) or independent samples (Y – O) t-tests. Numbers in boldface indicate significant results.

*Definition: R&K: lights out to 3 consecutive epochs S1 or the first epoch of any deeper sleep.

AASM: lights out to first epoch of any sleep.

min ($P = 0.000$; range: -8.9 to -30.8 min) and time in S3+S4/N3 decreased significantly by 9.1 min ($P = 0.006$; range: 4.7 to 10.3 min). In contrast, time in stage REM/R did not differ significantly by -1.8 min (NS; range: 0.1 to -1.8 min). Concerning the percent of TST in each stage, S1/N1 increased significantly by 2.8% ($P = 0.000$; range: 1.2% to 3.7%), S2/N2 decreased significantly by -4.9% ($P = 0.000$; range: -1.9% to -7.0%), S3+S4/N3 increased significantly by 2.4% ($P = 0.005$; range: 1.1% to 2.9%), and REM/R did not differ significantly by -0.4% (NS; range: 0.4% to -0.4%).

All tests for differences of the change values “AASM – R&K” between females and males were nonsignificant.

Young and Older Healthy Subjects

To determine whether the observed changes were age-dependent, a separate comparison was performed for younger and older healthy subjects, splitting the data into 2 groups of almost the same size, using an age cut-off of 60 years (see Table 1 for demographic data). As seen in Table 3, results for older subjects were similar to those of the total group. Accordingly, the ANOVAs revealed significant effects of the repeated measurement factor “classification standard” for WASO ($F = 9.27$, $df = 1,52$; $P = 0.004$), sleep stage N1 ($F = 16.88$, $df = 1,52$; $P = 0.000$ and $F = 17.75$, $df = 1,52$; $P = 0.000$ for min. and %TST, respectively), sleep stage N2 ($F = 23.15$, $df = 1,52$; $P = 0.000$ and $F = 22.28$,

$df = 1,52$; $P = 0.000$ for min. and %TST, respectively) and sleep stage N3 ($F = 8.23$, $df = 1,52$; $P = 0.006$ and $F = 8.23$, $df = 1,52$; $P = 0.006$ for min. and %TST, respectively). Interestingly, the time spent in stage REM was longer for the data based on R&K scorings than on AASM scorings in the young subjects. Indeed, the only significant interaction in the ANOVAs (“classification standard” × sex, “classification standard” × age, and “classification standard” × sex × age) was “classification standard” × age for stage R in minutes and in percent of TST ($F = 5.84$, $df = 1,52$; $P = 0.019$ and $F = 5.25$, $df = 1,52$; $P = 0.026$ for stage R in min. and %TST, respectively). All other interactions were nonsignificant. Accordingly, the only significant effects on the change values “AASM – R&K” between young and older subjects were seen for stage R in minutes and %TST (see last column of Table 2). All tests for differences of the change values “AASM – R&K” between females and males were nonsignificant.

These findings were confirmed by regression analysis with the change values “AASM – R&K” as dependent variable and age as independent variable. Only for the change values of stage R in min. and %TST a significant slope was found (0.16 min/year, $P = 0.009$ and 0.04 %TST/year, $P = 0.009$). The equations were $-10.8 \text{ min} + 0.16 \text{ min} * \text{age}$ (years). Thus, stage R was reduced by 7.5 min for 20-year-old subjects, while no reduction was observed for subjects aged 60 years or older. Multiple regression analysis with age and age-square as independent variables revealed no evidence of nonlinearities.

Table 4—Sleep Scoring Parameters (Mean ± Standard Deviation) Derived from Sleep Classification According to R&K and AASM for Patients (PAT, n = 16) as Well as for Age- and Sex-Matched Healthy Subjects (HC, n = 16)

	PAT: Patients				HC: Age- and sex-matched healthy subjects				PAT - HC	
	R&K	AASM	Diff	P	R&K	AASM	Diff	P	Diff	P
Lights out clock time (hr:min)	22:51 ± 00:37	22:51 ± 00:37	-	-	20:11 ± 07:53	20:11 ± 07:53	-	-	-	-
Lights on clock time (hr:min)	06:56 ± 00:41	06:56 ± 00:41	-	-	06:22 ± 02:32	06:22 ± 02:32	-	-	-	-
Total sleep time (TST; in min)	381.0 ± 66.5	380.3 ± 68.3	-0.7	0.875	411.7 ± 59.2	411.2 ± 57.6	-0.5	0.979	-0.2	0.915
Total recording time ("lights out" to "lights on" in min)	490.3 ± 51.6	487.0 ± 47.0	-3.3	0.403	491.1 ± 34.6	491.4 ± 34.8	0.3	0.427	-2.2	0.185
Sleep latency (SL; in min)*	14.7 ± 10.8	16.5 ± 17.4	1.8	0.479	16.4 ± 17.3	14.8 ± 18.8	-1.6	0.266	-3.6	0.433
Sleep REM/R latency (sleep onset to first epoch of Stage R in min)	99.5 ± 61.5	97.4 ± 43.4	-2.2	0.861	81.8 ± 30.9	81.8 ± 31.1	0.0	0.362	-2.2	0.626
Wake after sleep onset (WASO; Stage W after sleep onset, in min)	89.4 ± 53.1	90.1 ± 51.7	0.8	0.698	60.3 ± 37.6	64.7 ± 37.5	4.4	0.365	-3.6	0.779
Percent sleep efficiency (TST/total recording time)x100 in %	78.5 ± 11.5	78.2 ± 12.2	-0.3	0.778	83.8 ± 9.4	83.7 ± 9.4	-0.1	0.737	-0.2	0.484
Time in each stage (min)										
Stage S1/N1	47.0 ± 23.0	60.4 ± 28.5	13.4	0.019	48.2 ± 24.1	60.3 ± 26.5	12.1	0.011	1.2	0.133
Stage S2/N2	205.8 ± 55.0	192.4 ± 52.1	-13.3	0.163	229.2 ± 38.5	203.2 ± 44.7	-26.0	0.014	12.6	0.196
Stage S3+S4/N3	51.3 ± 27.0	54.1 ± 35.8	2.8	0.665	56.8 ± 38.6	70.8 ± 46.0	14.0	0.075	-11.3	0.460
Stage REM/R	76.9 ± 25.8	73.5 ± 27.1	-3.5	0.140	77.6 ± 27.0	76.8 ± 28.0	-0.8	0.736	-2.7	0.712
Percent of TST in each stage (time in each stage/TST)x100 in %										
Stage S1/N1	12.2 ± 5.2	15.8 ± 7.2	3.6	0.017	12.2 ± 6.5	15.2 ± 7.8	3.0	0.010	0.5	0.153
Stage S2/N2	53.4 ± 7.2	50.1 ± 8.4	-3.3	0.172	55.9 ± 7.1	49.5 ± 9.6	-6.4	0.011	3.1	0.254
Stage S3+S4/N3	14.2 ± 9.4	15.0 ± 12.2	0.8	0.642	13.4 ± 8.5	16.8 ± 10.2	3.4	0.064	-2.6	0.680
Stage REM/R	20.2 ± 6.1	19.1 ± 5.9	-1.1	0.136	18.5 ± 4.9	18.4 ± 5.4	-0.1	0.847	-1.0	0.941

The P-values are based on Wilcoxon matched pairs signed-ranks test (PAT, HC) or Mann-Whitney U-test (PAT – HC) for TST, SL, REM/R-latency, WASO and sleep efficiency, for all other variables on paired samples (PAT, HC) or independent samples (PAT – HC) *t*-tests. Numbers in boldface indicate significant results.

*Definition: R&K: lights out to 3 consecutive epochs S1 or the first epoch of any deeper sleep.

AASM: lights out to first epoch of any sleep.

To explore age-related changes between stage REM and stage R in greater detail, the number of REM phases as well as their average duration were determined according to the definition of sleep cycles by Feinberg and Floyd¹¹ and regression analyses were performed for the change values of these variables as well. Interestingly, no age-related changes were observed for the number or for the duration of the REM phases. Thus, the only explanation for reduced time in stage R without a change in the duration of phase R is an increased time with NREM intrusions in phase R for young subjects (see also Figure 4 for example).

Patients Versus Healthy Subjects

To determine whether the observed changes in sleep parameters were different for healthy subjects and patients, the results for the 2 groups are contrasted in Table 4. The analyzed patient sample consisted of 5 patients with general anxiety disorder (GAD), 5 patients with periodic limb movement disorders (PLMD), and 6 patients with Parkinson disease. For the purpose of comparison, a group of 16 age- and sex-matched controls was derived from the sample of healthy subjects. The effect of the sleep scoring standard was more pronounced in healthy subjects than in patients. ANOVAs revealed no significant interactions between the factors "classification standard" and diagnosis. Moreover, the test for differences in the change

values "AASM – R&K" between patients and controls was nonsignificant. However, due to the relatively small number of patients as well as the non-homogeneity of the sleep disorders, results have to be interpreted with care.

Single Cases

In addition to the results for group statistics, 4 single cases are shown, which provide typical examples of the effects of the scoring standard. Figures 1 to 4 depict the hypnograms for the sleep classification according to R&K and AASM in the upper part and a typical 30-s epoch of raw data in the lower part of the figures.

Case 1: No Change

In the first example of a 24-year-old healthy man no relevant changes can be seen between the 2 different visual analyses, neither in the hypnogram (Figure 1) nor in the derived sleep parameters. In the 30-s epoch displayed in the lower part of Figure 1, high slow wave amplitudes at both central and frontal derivations (200 μ V and more) can be seen. Note the relatively small number of cortical arousals and the high delta and spindle density as identified by the Somnolyzer 24x7¹² adapted for the automatic classification according to the AASM criteria (upper part of Figure 1). Sleep stage scoring by the experts was per-

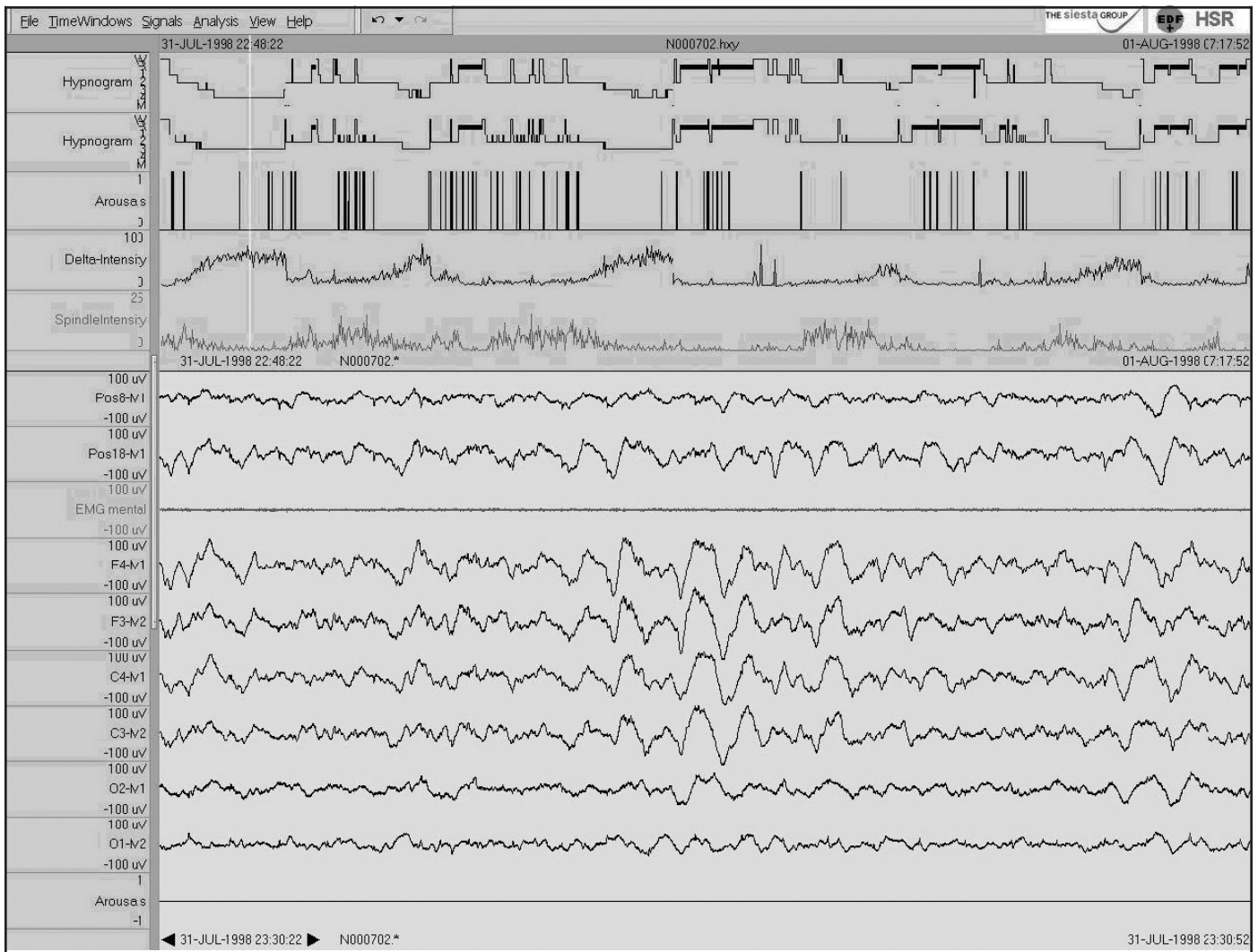


Figure 1—Single case 1: 24-year-old man (healthy subject) who shows “no changes” in the sleep scoring parameters. In addition to the R&K and AASM hypnograms (first and second traces, respectively), arousals, delta and spindle intensity as revealed by the Somnolyzer 24x7 adapted for the AASM standard are shown in the upper part of the figure. The lower part shows a typical 30-s epoch in sleep stage N3 with 2 EOG derivations (Pos8-M1 and Pos18-M1), a mental EMG derivation, and 6 EEG derivations (F4-M1, F3-M2, C4-M1, C3-M2, O2-M1, O1-M2), as well as a channel indicating automatically detected arousals. To facilitate comparisons, all scales are the same for Figures 1 to 3. Note that the high-amplitude slow waves surpass the 75 μV criteria by far already at central leads.

formed on the basis of raw data only. The automatically classification shown in Figure 1 was not presented.

Case 2: S3+S4/N3 Increase

In the second example of a 32-year-old healthy woman, the major change from the R&K to the AASM scoring is an increase in slow wave sleep from 105.5 to 139.5 min (22.6% to 29.6% of TST) at the cost of a decrease in stage S2/N2. In many epochs the amplitudes of the slow waves are just below 75 μV at central leads, but clearly above 75 μV for frontal leads. For details see Figure 2.

Case 3: S1/N1 Increase

In this study of a healthy 80-year-old man, an increase in light sleep from 125.0 to 141.0 min (31.8% to 36.2% of TST) is seen for AASM compared to R&K scoring (Figure 3). Because

the subject showed only few sleep spindles and K-complexes as well as a high number of arousals (see upper part of Figure 3), many 30-s epochs scored as S2 according to R&K changed to N1 according to AASM, since an arousal defines the end of stage N2 sleep until a sleep spindle or a K-complex unassociated with an arousal occurs (rule 5.C.1.b in the AASM manual⁶). For an example see lower part of Figure 3. Visual scorings were based on raw-data only; see above.

Case 4: REM/R Decrease

In the fourth example of a 22-year-old healthy woman, a decrease in stage R sleep from 101.0 to 79.5 min (21.6% to 18.3% of TST) is seen for AASM compared to R&K scoring (Figure 4). Interestingly, the decrease in stage R is predominantly due to an increased time with NREM intrusions, specifically in the first and second REM phases. It may be that the additional information from frontal (and occipital) leads led to

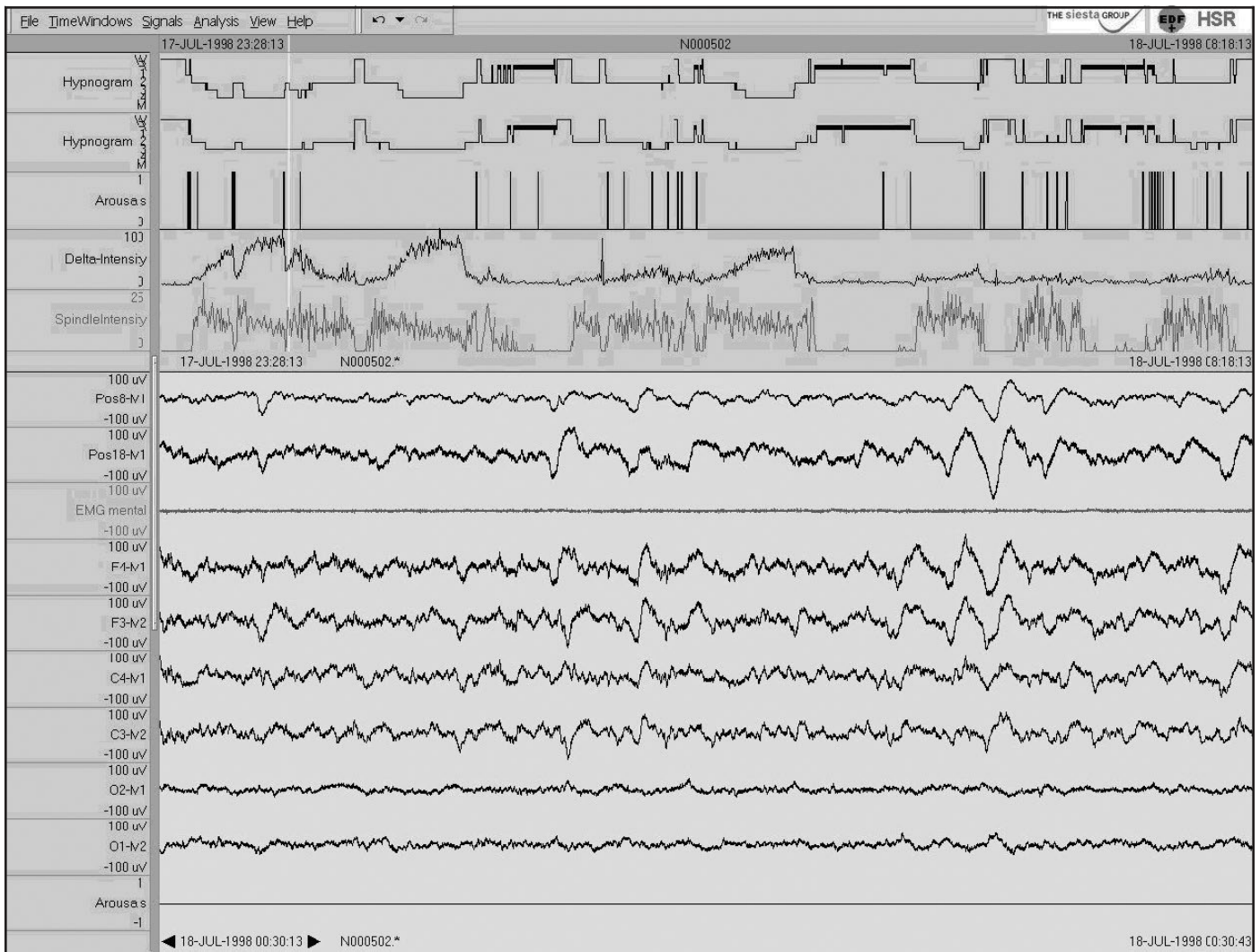


Figure 2—Single case 2: 32-year-old woman (healthy subject) who shows an “increase in N3/SWS” in the sleep scoring parameters. For details see Fig. 1. Note that the amplitudes of the slow waves at central leads are sometimes just below $75 \mu\text{V}$, while the amplitude criterion of $75 \mu\text{V}$ is reached at frontal leads.

longer and more frequent NREM intrusions—probably due to a better identification of (small-amplitude) K-complexes and/or spindles (see lower part of Figure 4).

DISCUSSION

The aim of the present study was to evaluate differences between visual sleep scorings according to the R&K¹ classification and scorings based on the new AASM⁶ standards. Both scoring standards were applied by experienced sleep experts from different European sleep labs. In order to ensure that scorings according to the new AASM standard were performed correctly, a 2-day training symposium was organized by The Siesta Group in Vienna. During this course the AASM scoring manual as well as the slides from course C11: “Advanced Polysomnography 2007: Understanding and Using the New AASM Sleep Scoring Manual” held at the 21st Annual Meeting of the APSS “Sleep 2007” in Minneapolis were discussed in detail, and both combined and individual scorings of 4 PSGs were performed. The results of the individual scorings were compared and discussed to explain and solve possible incongruence. In a separate paper by

Danker-Hopfe et al,¹³ the results for interrater reliabilities (IRRs) between the expert scorings according to the AASM standard are presented. The authors compared the IRRs obtained for AASM scorings with those obtained for R&K scorings using intraclass correlations for the derived sleep parameters as well as Cohen’s and Fleiss’ kappa for epoch-by-epoch comparisons. The results showed that IRR was higher for scorings according to the AASM standard for all stages, except S2/N2. The results of the present study are based on the R&K consensus scoring (after 2 independent scorings) obtained from the SIESTA project and on the AASM scoring of well-trained sleep experts (first scorers) whose performance was checked using the scoring of a second expert. In addition, the minimal and maximal differences obtained in all possible pair-wise comparisons of AASM 1st and 2nd scorings with R&K 1st and 2nd scorings are given to estimate the effect of IRR on the magnitude of the observed changes.

With respect to the visual classification of sleep, the new AASM standard extends the R&K standard at 3 levels: (1) data acquisition, (2) scoring rules, and (3) definition of sleep scoring parameters. In the following paragraphs, the effect of these changes on the sleep scoring parameters will be discussed.

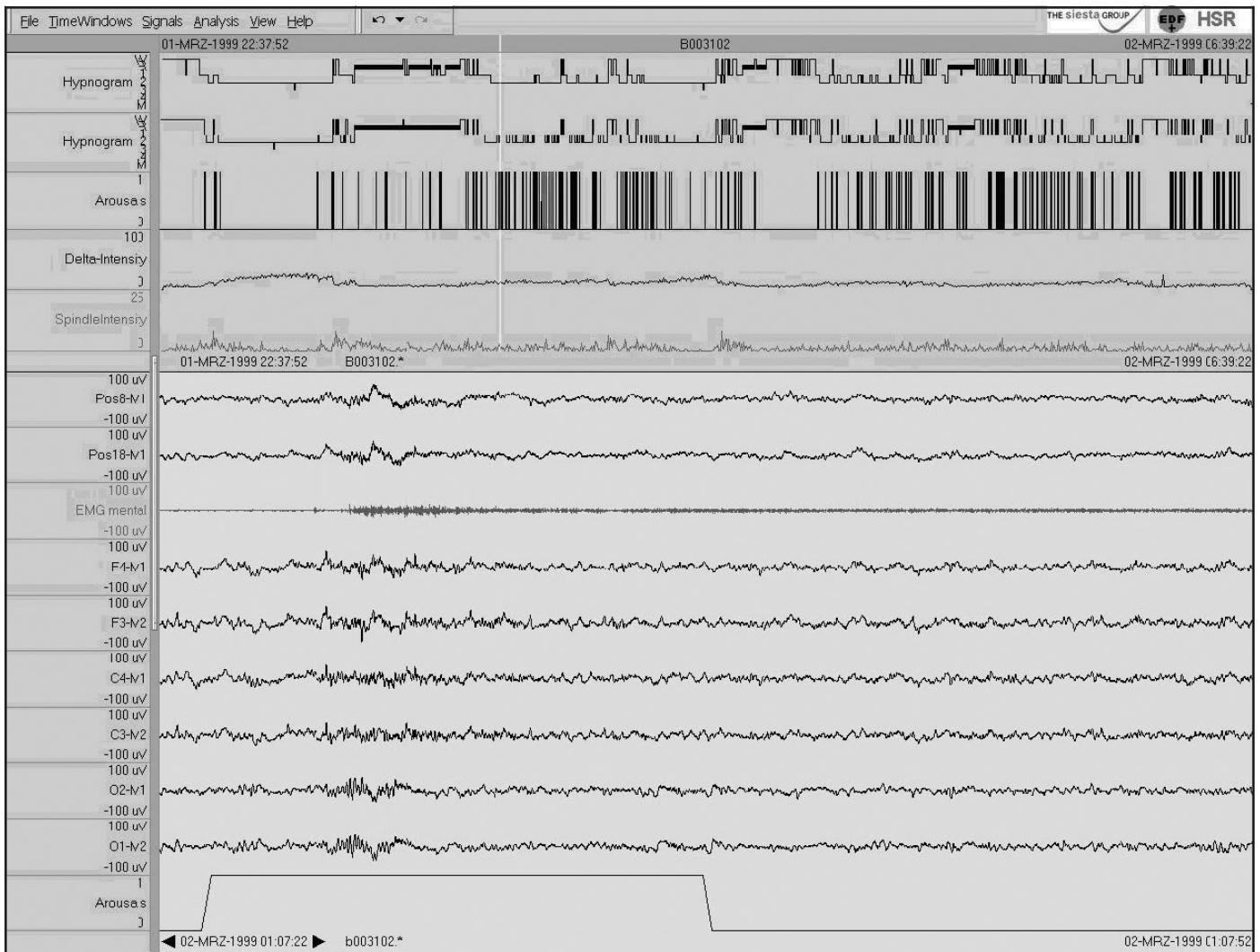


Figure 3—Single case 3: 80-year-old-man (healthy subject) who shows an “increase in N1/S1” in the sleep scoring parameters. For details see Fig. 1. Note the relatively high number of arousals and the low delta and spindle intensity. As seen specifically in the second sleep cycle, many sleep stages scored as S2 according to R&K were scored as N1 according to AASM because of cortical arousals. For an example see the lower part of the figure.

While the definition of the specifications for data acquisition, such as minimal sampling rates, filter settings and display resolutions, is highly appreciated, most of these recommendations have been standard in many sleep labs for years and thus should not affect study results significantly. Although the present data were recorded between 1998 and 1999, all recordings fulfilled these minimal specifications already. Moreover, the additional recording of frontal and occipital EEG derivations was already recommended by the R&K committee in 1968 (see figures 33–40 in the R&K manual⁵). In the new standard, however, derivations from frontal and occipital regions are obligatory and therefore information from these channels was included in the scoring rules.

Most of the observed changes in sleep scoring parameters from the R&K to the AASM standard are certainly due to modifications of the scoring rules. The inclusion of frontal derivations for the identification of slow wave activity led to a significant increase in slow wave sleep (approximately + 10 min or 2.5% of TST). This increase is in accordance with topographic sleep EEG studies showing the maximal amplitudes of slow

waves at frontal leads.^{14,15} This increase, however, was specifically marked in subjects with borderline slow wave amplitudes at central derivations (see Figure 2). For subjects with high slow wave amplitudes at central leads, no significant increase from stage S3+S4 to N3 was found (Figure 1). But for subjects with low amplitude 0.5 Hz–2 Hz waves at central leads, only small increases in stage N3 were observed, since for these subjects amplitude of low-frequency activity was below the threshold of 75 μ V at frontal derivations as well (Figure 3).

Epochs that changed to N3 had previously been scored S2: approximately half of the 20-min decrease from S2 to N2 is explained by the inclusion of frontal leads in slow wave detection. The other half reflects a change from S2 to N1, explained by the new scoring rule 5.C.b (see page 26 of the AASM manual⁶), which states that a cortical arousals defines the end of N2. The AASM task force stated in the paper accompanying the AASM manual⁷ that they intended to consider situations in which sleep is highly fragmented by arousals. While an arousal not associated with a “pronounced increase in muscle tone” does not indicate the end of stage S2 according to R&K, stage N2 ends

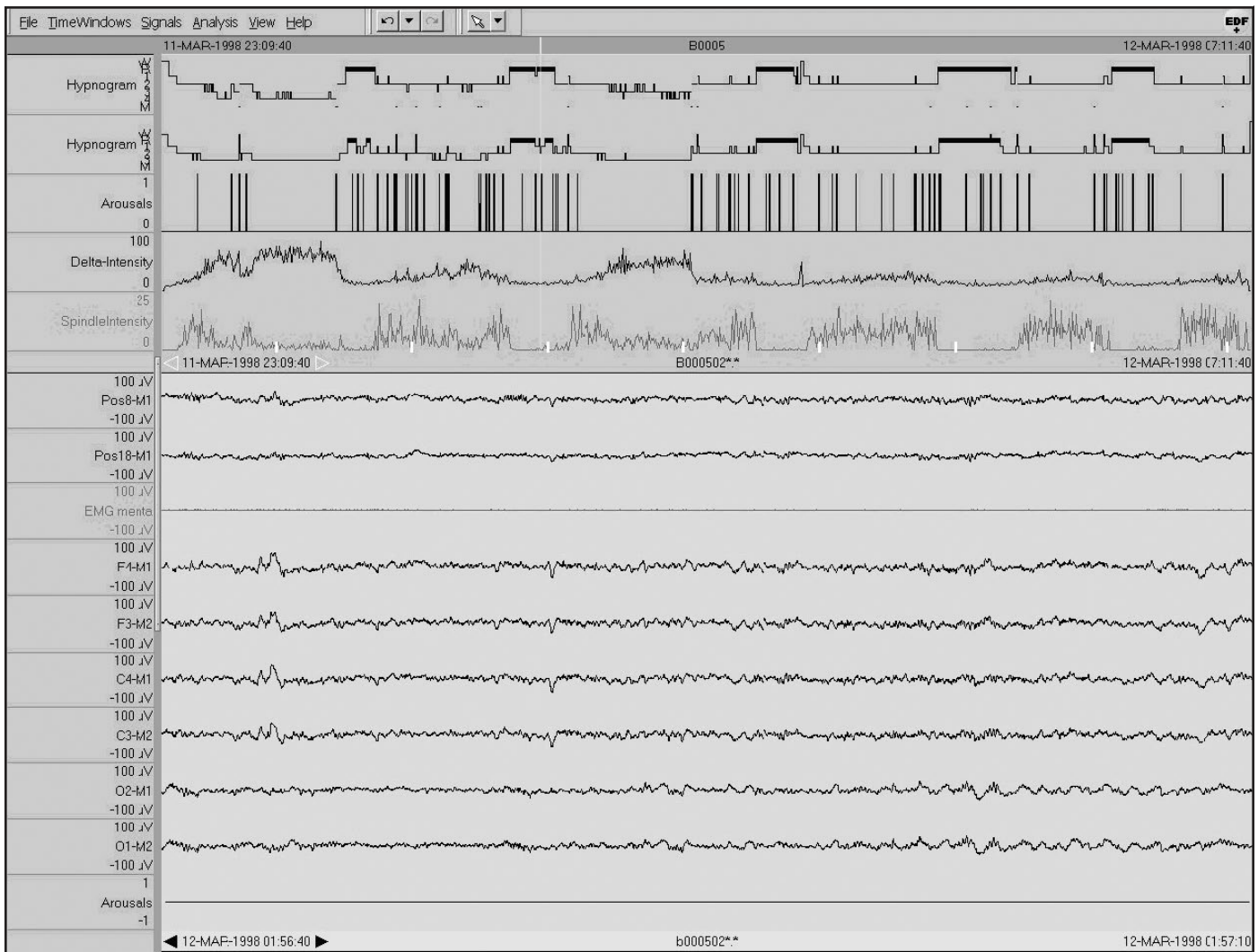


Figure 4—Single case 4: 22-year-old woman (healthy subject) who shows a “decrease in REM/R.” For details see Figure 1. Note the longer NREM intrusions in the first and second REM phase in the AASM hypnogram. The relatively low-amplitude K-complex in the central leads at the beginning of the 30-s example is confirmed by its typical topographic distribution with a fronto-central maximum (see lower part of the Figure).

with an arousal irrespective of a concurrent increase in EMG tone. Since after an arousal in NREM sleep, N2 should not be scored again until a K-complex unassociated with an arousal or a sleep spindle occurs, the extent of change from S2 to N1 depends on both the arousal and the spindle and K-complex indices. The higher the arousal index and the lower the spindle and K-complex indices, the more epochs will change from S2 to N1. Since the number of arousals increases with advancing age¹⁶ and the number of spindles and K-complexes decreases,¹⁷ it is likely that this effect will be more pronounced in older subjects (see case 3).

Interestingly, the “3-min rule” for the maximal length of a period without spindles or K-complexes in S2 had only minor effects in the present study. To estimate the influence of omitting the “3-min rule” we performed a computerized analysis according to the new AASM standard (Somnolyzer) and compared the minutes in N2 with a modified software version that changed to N1 after a 3-min period, simulating the R&K “3-min rule.” The decrease in N2 due to this modification was approximately 1.5 min on average. Thus, the “3-min rule” rarely changed the scor-

ing of N2. Indeed, this result confirms that 3 min for this rule were well chosen by the R&K committee. In summary, while 1-2 min of S1 change to N2 due to the omission of the “3-min rule,” more than 10 min of S1 change to N2 due to the inclusion of the “arousal rule.”

Concerning stage REM, the AASM rules adhere relatively closely to the R&K standard and thus, only minor changes from stage REM to stage R were observed in the present study. The 4.5-min decrease in stage R in the group of younger healthy subjects was not seen in any other group comparisons. Interestingly, however, this decrease in stage R is not due to a shortening of the R phase within the sleep cycles, but rather to more frequent and longer NREM intrusions, probably initiated by the additional information from frontal and occipital EEG leads for the identification of NREM-related sleep patterns such as K-complexes and spindles (Figure 4 for an example). Since older subjects have reduced K-complex and spindle densities, these intrusions specifically affect stage R in younger subjects. The major advantage of the new rules for scoring REM is the more systematic description for scoring this stage.

The observed increase in wake after sleep onset (WASO) from the R&K to the AASM rules (approximately + 4 min) is partly due to the change in the rule for scoring movement time and, to a larger extent, to a change in the definition of sleep onset. In the present study, an average of 2 of 3 epochs scored as movement time were scored as stage W, since alpha rhythm was detectable in some parts of these epochs. Sleep onset has been defined by 3 consecutive stages of S1 or any other deeper sleep stage; this definition was used for the determination of WASO based on R&K scoring.¹⁸ According to the AASM manual, however, sleep onset is defined as the first epoch of any sleep stage; as a consequence, the number of epochs scored as W after sleep onset may increase. Interestingly, while for the total group WASO increased significantly by 4 min (65.9 min and 69.9 min for WASO according to R&K and AASM, respectively), this increase was only 1.8 min and not significant if for both the AASM and the R&K scorings sleep onset was defined as the first epoch of sleep (68.1 min and 69.9 min for R&K and AASM scorings, respectively). Since most of the increase in WASO for the AASM scorings is due to an advancement of sleep onset and not to a change from sleep to wakefulness, neither total sleep time nor sleep efficiency changed significantly from the R&K to the AASM standard. The lack of significance in the comparison of sleep onset latencies might be due to the relatively high variance in the change values between AASM and R&K for this measure. In 29 recordings (40%), no differences were found between sleep latencies defined by one epoch of any sleep stage and sleep latencies defined by 3 consecutive light sleep stages or any deeper sleep stage. In 15 recordings (21%), the difference was 0.5 min to 1 min. In 10% of the recordings, however, the difference was 5 min or more and the maximal difference was 19.5 min.

In conclusion, the new standard showed only a minor influence on total sleep time, sleep efficiency, and stage REM, but affected WASO and the distribution of NREM sleep stages. Studies establishing normative data for sleep have so far been based on sleep classification according to the rules of R&K¹⁹⁻²¹ (for a review and meta-analysis see Ohayon²²). Since the present study has shown numerically relevant differences between sleep parameters derived AASM and R&K rules, it is no longer possible to use the existing age-dependent normative values for these parameters; new sleep studies in large and representative control groups must be conducted to fill this gap. One challenge for a sleep expert scoring sleep visually according to the new AASM standard is the identification of cortical arousals in NREM, specifically those not associated with increased EMG tonus. The actual implications for an individual study depend on slow wave amplitudes and their anterior-posterior gradient as well as on the density of arousals, spindles, and K-complexes and may range from no changes at all to increases of 53 min from S3+S4 to N3 and 44 min from S1 to N1, with a consequent decrease from S2 to N2 of 55 min. As revealed by pair-wise comparisons between independent scorings, the range for mean increases from S3+S4 to N3 was between 5 and 10 min; from S1 to N1, it was between 3 and 13 min. The mean decreases from S2 to N2 ranged between 9 and 31 min. Finally, the increases in WASO ranged between 2 and 10 min. In summary, even though IRR certainly influences the exact figures of the reported differences, shifts in NREM sleep from N2 to N1 and N3 and increase in WASO were consistently observed.

Finally, the new standard has no direct influence on diagnosis according to ICSD-2 because its diagnostic criteria are not primarily based on polysomnographic features. Some recommendations (e.g., psychophysiological insomnia, idiopathic insomnia, narcolepsy) include information on sleep latency, but no quantitative data on the distribution of sleep stages.

ACKNOWLEDGMENTS

This study was supported by The Siesta Group Schlafanalyse GmbH, Vienna. The authors would like to express their thanks to Mag. Elisabeth Grätzhofer for her valuable editorial assistance.

DISCLOSURE STATEMENT

This study was supported by The Siesta Group Schlafanalyse GmbH, Vienna, Austria. Dr. Anderer is a shareholder and Chief Science Officer of the Siesta Group. Mr. Gruber, Dr. Parapatics, and Ms. Loretz are employees of The Siesta Group Schlafanalyse GmbH. Dr. Danker-Hopfe is a shareholder and is supervisory board chairman of The Siesta Group Schlafanalyse GmbH. Dr. Saletu has received research support from Abiogen Pharma, Actelion, Astra Zeneca, Cephalon, GlaxoSmithKline, Schwarz Pharma, Servier, and Takeda; has been an advisory board member for Nycomed, Servier, Takeda, and UCB; has consulted for Merck and Xenoport; has participated in speaking engagements for Astra Zeneca, Cephalon, Ixico, Janssen, and Lundbeck; and is a shareholder of The Siesta Group Schlafanalyse GmbH. Dr. Zeitlhofer is a shareholder of The Siesta Group Schlafanalyse GmbH. Dr. Dorffner is CEO, Managing Director and 10% owner of The Siesta Group Schlafanalyse GmbH. The other authors have indicated no financial conflicts of interest.

REFERENCES

1. Rechtschaffen A, Kales A, eds. A manual of standardized terminology, techniques and scoring system of sleep stages in human subjects. Los Angeles: Brain Information Service/Brain Research Institute, University of California, 1968.
2. Penzel T, Behler PG, Buttler von M, et al. Reliabilität der visuellen Schlafauswertung nach Rechtschaffen und Kales von acht Aufzeichnungen durch neun Schlaflabore. *Somnologie* 2003;7:49-58.
3. Danker-Hopfe H, Kunz D, Gruber G, et al. Interrater reliability between scorers from eight European sleep laboratories in subjects with different sleep disorders. *J Sleep Res* 2004;13:63-9.
4. Kubicki St, Herrmann WM, Höller L, Scheuler W. Kritische Bemerkungen zu den Regeln von Rechtschaffen und Kales über die visuelle Auswertung von EEG-Schlafableitungen. *EEG – EMG* 1982;13:51-60.
5. Himanen SR, Hasan J. Limitations of Rechtschaffen and Kales. *Sleep Med Rev* 2000;4:149-67.
6. Iber C, Ancoli-Israel S, Chesson A, Quan SF, eds. The AASM manual for the scoring of sleep and associated events: rules, terminology, and technical specification, 1st ed. Westchester, IL: American Academy of Sleep Medicine, 2007.
7. Silber MH, Ancoli-Israel S, Bonnet MH, et al. The visual scoring of sleep in adults. *J Clin Sleep Med* 2007;3:121-31.
8. Kloesch G, Kemp B, Penzel T, et al. The SIESTA-Project – polygraphic and clinical database: a new approach to studying subjective and objective measurements of human sleep. *IEEE* 2001;20:51-7.

9. Penzel T, Kemp B, Kloesch G, et al. Acquisition of biomedical signals databases: aspects to consider when building a database, based on experiences from the SIESTA Project. *IEEE* 2001;20:25-32.
10. Buysse DJ, Reynolds III CHF, Monks TH, Berman S, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatr Res* 1989;28:193-213.
11. Feinberg I, Floyd TC. Systematic trends across the night in human sleep cycles. *Psychophysiology* 1979;163:283-91.
12. Anderer P, Gruber G, Parapatics S, et al. An e-health solution for automatic sleep classification according to Rechtschaffen and Kales: validation study of the Somnolyzer 24x7 utilizing the Siesta database. *Neuropsychobiology* 2005;51:115-33.
13. Danker-Hopfe H, Anderer P, Zeitlhofer J, et al. Interrater reliability (IRR) for sleep scoring according to the Rechtschaffen & Kales and the new AASM standard. *J Sleep Res* 2008 (in press).
14. Werth E, Achermann P, Borbely AA. Fronto-occipital EEG power gradients in human sleep. *J Sleep Res* 1997;6:102-12.
15. Happe S, Anderer P, Gruber G, Klosch G, Saletu B, Zeitlhofer J. Scalp topography of the spontaneous K-complex and of delta-waves in human sleep. *Brain Topogr* 2002;15:43-9.
16. Bonnet MH, Arand DL. EEG arousal norms by age. *J Clin Sleep Med* 2007;3:271-4.
17. Crowley K, Trinder J, Kim Y, Carrington M, Colrain IM. The effects of normal aging on sleep spindle and K-complex production. *Clin Neurophysiol* 2002;113:1615-22.
18. Carskadon MA, Rechtschaffen A. Monitoring and staging human sleep. In: Kryger MK, Roth T, Dement WC, eds. *Principles and practice of sleep medicine*. 4th ed. Philadelphia: Elsevier 2005:1359-77.
19. Hume KI, Van F, Watson A. A field study of age and gender differences in habitual adult sleep. *J Sleep Res* 1998;7:85-94.
20. Walsleben JA, Kapur VK, Newman AB, et al. Sleep and reported daytime sleepiness in normal subjects: the Sleep Heart Health Study. *Sleep* 2004;27:293-8.
21. Danker-Hopfe H, Schaefer M, Dorn H, et al. Percentile reference charts for selected sleep parameters for 20- to 80-year-old healthy subjects from the SIESTA database. *Somnologie* 2005;9:3-14.
22. Ohayon MM, Carskadon MA, Guilleminault Ch, Vitiello MV. Meta-Analysis of quantitative sleep parameters from childhood to old age in healthy individuals: developing normative sleep values across the human lifespan. *Sleep* 2004;27:1255-73.