The AASM Scoring Manual Four Years Later

Madeleine M. Grigg-Damberger, M.D., F.A.A.S.M.

University of New Mexico, Albuquerque, NM

Purpose of Review: Review published studies and critiques which evaluate the impact and effects of the American Academy of Sleep Medicine (AASM) Sleep Scoring Manual in the four years since its publication.

Findings: Using the AASM Manual rules to score sleep and events in a polysomnogram (PSG) results in: (1) very large differences in apnea-hypopnea indexes (AHI) when using the recommended and alternative rule for scoring hypopneas in adults; (2) increases in NREM 1 and sleep stage shifts with compensatory decreases in NREM 2 in children and adults when following rule 5.C.b. for ending NREM 2 sleep; (3) increases in NREM 3 in adults scoring slow wave activity in the frontal EEG derivations; (4) improved interscorer reliability; and (5) successfully identified fragmented sleep in children with obstructive sleep apnea (OSA) from primary snorers or normal controls because they had more NREM 1 and stage shifts using rule 5.C.b. Criticism of the Manual most often cited: (1) two rules for scoring hypopneas; (2) alternative EEG montage cancellation effects; (3) scoring stages 3 and 4 as NREM 3; and (4) too few rules for scoring arousals and REM sleep without atonia.

Summary: Four years have passed since the AASM Scoring Manual was published with far less criticism than those who developed it feared. The AASM Manual provides a foundation upon which we all can build rules and methods which identify the complexity of sleep and its disorders.

Keywords: Sleep scoring parameters, AASM standard, polysomnography/classification/standards, practice guidelines as topic sleep apnea, practice guidelines as topic sleep scoring

Citation: Grigg-Damberger MM. The AASM Scoring Manual four years later. *J Clin Sleep Med* 2012;8(3):323-332.

Four years have passed since the American Academy of Sleep Medicine (AASM) published *The AASM Manual for the Scoring of Sleep and Associated Events*.[1] The Manual represents a brave effort to standardize how a comprehensive in-laboratory polysomnogram (PSG) performed in the United States should be recorded, scored, analyzed, and reported.[2] Before it, we were never certain that sleep study data across city, state, country, or sea were comparable. Based on digital video-PSG recording techniques, the Manual incorporates the effects of age and disease on sleep and provides rules for visual scoring of sleep stages, arousals, movements, respiratory and cardiac events during sleep.[2,3] Driving goals in developing it were: simplicity, ease of implementation, likelihood of improving interscorer reliability, and avoiding radical changes unless sufficient evidence existed to do so. Every effort was made to make the Manual evidence-based, but when insufficient, the committee used the UCLA/Appropriateness Method to develop consensus recommendations.[4] This paper reviews studies and critiques which evaluate the impact and effects of the AASM Manual Sleep Scoring Manual in the four years since its publication.

Two Definitions for Scoring Hypopneas in Adults

The greatest criticism has been directed at the AASM's decision to permit two different ways for scoring hypopneas in a PSG in adults.[1] Why two? The Respiratory Scoring Task Force detailed in its review paper[5] convincing level I and II evidence that apnea-hypopnea indexes (AHI) which define hypopneas as only a discernible amplitude reduction but were associated with a $\geq 4\%$ desaturation correlated significantly with: (1) prevalent and incident blood pressure;[6,7] (2) lower neurocognitive test scores;[8] and (3) automobile crashes due to sleepiness.[9] The respiratory review paper also provided sufficient evidence for scoring hypopneas which caused as little as $\geq 20\%$ to 30% fall in airflow and/or $\geq 2\%$ or $\geq 3\%$ desaturations.[10-12]

However, the AASM wanted rules for scoring hypopneas which had substantial or excellent interscorer reliability. Inter-rater reliability (IRR) as it relates to scoring sleep in a PSG measures how closely individuals score the same sleep study. IRR when scoring a PSG depends upon the: (1) skill, experience, and training of scorer; (2) technical quality of study; (3) clarity and simplicity of scoring rules; (4) diligence with which scoring rules are applied; and (5) degree of physiological ambiguity of the sleep/wake patterns.[13] When two or more individuals score a stage of sleep or an event in a PSG differently, it can introduce enough variability to lead to a false positive or false negative for a particular diagnosis.

The IRR for scoring respiratory events in a PSG is particularly affected by the: (1) duration of the event; (2) degree of reduction in the amplitude of the measured signal(s); (3) level of the oxyhemoglobin desaturation associated with it; and (4) presence and duration of arousal which accompanies it.[5]

Ayappa et al. found the percent scoring agreement for scoring apneas using a nasal pressure sensor was excellent (0.91) but only moderate for hypopneas (0.69) or flow limitation events (0.64).[14] The more subtle the reduction in airflow, the more difficult to achieve good interscorer agreement: one scorer identified 35% more flow limitation events than the other. Whitney et al. found a hypopnea associated with 2% to 5% desaturations could be scored with IRR of 0.90, but scoring agreement in-

creased only 0.04 to 0.94 if an arousal accompanied the hypopnea.[15] Scoring agreement for hypopneas which did not cause an arousal or desaturation was only 0.74, and 0.77 for those which only caused an arousal.

After reviewing this evidence, the respiratory task force proposed that hypopneas in adults were most likely to be scored with excellent IRR (IRR $\geq$ 0.80) if associated with $\geq$ 50% fall in airflow and $\geq$ 3% desaturation or $\geq$ 3-sec EEG arousal. However, concerns were expressed that this more "liberal" hypopnea definition might not be readily accepted by the Centers for Medicare and Medicaid Services (CMS). It had taken the AASM years to convince CMS that a hypopnea was as pathologic as an apnea. Hypopneas which caused $\geq$ 30% reduction in nasal pressure accompanied by $\geq$ 4% desaturation could be used to calculate an AHI to obtain financial support for a positive airway pressure therapy to treat adult OSA. To avoid a reimbursement hiatus and billing chaos, the committee decided to offer two hypopnea definitions. They proposed that the 50%/3%/AR rule be the recommended criteria for scoring hypopneas in an adult, with the 30%/4% an alternative scoring choice. However, the AASM Board of Directors chose the reverse, selecting the 30%/4% rule as the recommended rule   Several months after the Manual was published, the Manual Steering Committee advised that: (1) investigators use the alternative hypopnea (50%/3%/AR) rule in all prospective epidemiological and outcome studies; (2) for clinical purposes, sleep specialists may select either the recommended (page 46, 4.A) or alternative (page 46, 4.B) rule; and (3) for comparison purposes in clinical research or practice, both methods may be reported (FAQ, www.aasmnet.org).[16]

Impact of Two Hypopnea Definitions on the Apnea-Hypopnea Indexes

Ruehland et al. were first to evaluate how AHI and hypopnea indexes (HI) would be significantly different depending upon which AASM hypopnea definition was used to score an adult PSG.[17] They rescored 328 adult PSGs using the AASM Manual respiratory scoring criteria. These PSGs had originally been scored using hypopnea rules recommended by the AASM in 1999 (and often called the Chicago criteria).[18,19] The Chicago criteria permit scoring a hypopnea in an adult if there was a > 50% airflow reduction alone *or* a lesser airflow reduction associated with $\geq$ 3% desaturation or $\geq$ 3-sec EEG arousal (50% or less/3%/AR).

Ruehland et al. found very large and clinically significant differences in median AHI and hypopnea indexes (HI) depending upon how hypopneas were scored. The median AHI when they scored hypopneas using Chicago criteria (50% or less/3%/AR) was 25/h, 40% lower when they were scored using the alternative rule (AHI 15/h), and 70% lower with the recommended (AHI 8/h). Using the AASM-recommended 30%/4% hypopnea rule, 36%, 43%, and 48% of patients previously classified as positive for OSA using an AHI following Chicago (50% or less/3%/AR) criteria would now be negative with AHI cutoffs of 5, 15, and 30/h, respectively. Scoring hypopneas using AASM alternative 50%/3%/AR hypopnea rule, 17%, 26%, and 25% of patients previously classified as positive for OSA using Chicago criteria would now be negative for the diagnosis if the same benchmark AHI cutoffs were followed. The au-

thors cautioned cutoffs for abnormal AHI in a PSG to diagnose OSA need to be adjusted: an AHI of 5/h as scored using the recommended 30%/4% rule would correspond to 10/h using alternative 50%/3%/AR rule and 15/h using Chicago 50% or less/3%/AR criteria. Based upon any cut-point value of AHI, different hypopnea definitions would result in substantial and clinically significant differences for identifying and classifying severity of OSA. These findings prompted the authors to recommend the AASM endorse a single adult hypopnea definition.

Guilleminault et al. rescored 35 PSGs in lean adults with clinically symptomatic OSA who often had hypopneas which did not cause significant desaturations.[20] Mean AHI was 27/h when they scored these PSGs years earlier using the Chicago (50% or less/3%/AR) hypopnea scoring criteria.[18] Mean AHI using the AASM alternative 50%/3%/AR hypopnea definition was 22% lower (21/h), and 78% lower using the recommended 30%/4% rule (AHI 6/h). If PSGs were scored using recommended 30%/4% rule, PSGs in 14 (40%) would have been classified as normal (AHI < 5). They concluded that the AASM-recommended hypopnea rule should not be used to score hypopneas in lean subjects in whom many clinically significant hypopneas caused little/no desaturation.

Impact of Neglecting to Score Apneas or Hypopneas in Wake on Apnea-Hypopnea Indexes

Norman et al. studied the impact of not scoring apneas or hypopneas during epochs of stage W had on the AHI using computerized PSG scoring.[21] They first manually scored sleep in 60 diagnostic adult PSGs in 30-sec epochs; then rescored them using a method similar to adaptive segmentation.[22] They found 8% of apneas or hypopneas were not scored because they occurred within a single epoch of Wake (43%) or began in W but continued into sleep (57%). The median AHI increased from 15.8 to 18.8/h when they included respiratory events in stage W. The biggest discrepancies in AHI occurred in patients with AHI > 40/h (and which would not change the diagnosis or clinical management). Of note, the AASM Scoring Manual FAQ website states that a respiratory event can be scored if any part of it occurs during an epoch of sleep, so only 3% of the events in the study PSGs would *not* have been scored using AASM rules. The authors argue that absolute sleep scoring (a form of adaptive segmentation) would provide a truer measure of TST by counting short periods of sleep in a PSG often scored as stage W when scoring a PSG in 30-sec epochs. Lack of scoring respiratory events which cause or occur in W prevents including respiratory events that are contributing to sleep fragmentation and clinically important symptoms in patients with OSA.

Impact of Scoring Respiratory Events in Adolescents Using Pediatric or Adult Rules

To provide adolescents greater access to sleep laboratories, the AASM Manual permits scoring respiration in sleep in those $\geq$ 13 years using adult scoring criteria.[23] The AASM pediatric hypopnea rule is nearly identical to the adult alternative hypopnea (50%/3%/AR) rule except for event duration ($\geq$ 2 missed breaths in a child vs. $\geq$ 10 sec in an adult).[23] Accardo et al. retrospectively scored PSGs in 101 healthy non-obese adolescents (ages 13 to 18 years) without sleep/wake complaints.[24] The investigators found far lower mean AHI (0.4/h) when the

AASM-recommended 30%/4% hypopnea rule was used to score adolescent PSGs, compared to an AHI of 1.7/h using the pediatric respiratory scoring rules and 1.4/h using the alternative adult criteria. They found that far fewer adolescents would meet AHI ≥ 5 criteria for diagnosing OSA if their PSGs were scored using the adult recommended 30%/4% rule.

Lin and Guilleminault retrospectively scored PSGs in 209 non-obese children (2-18 years) with suspected OSA using Stanford and AASM pediatric respiratory criteria.[25] They found 99% (207/209) would be diagnosed as having OSA using Stanford criteria but only 19% (39/209) using AASM pediatric rules (AHI ≥ 1/h).[23,26] The median AHI of PSGs scored using Stanford criteria was 12.7/h, 0.1/h using the AASM. Based on their results, the authors asked the following questions and lodged these complaints regarding the AASM pediatric respiratory scoring criteria: (1) the AASM criteria identify only the most obvious and severe cases of pediatric OSA and lead to a high number of false-negatives; (2) why do a PSG if the scoring criteria cannot discriminate between patients with or without OSA symptoms?; (3) it is not acceptable for a test that is potentially demanding, costly, and used to make important health care decisions; and (4) if clinical symptoms and clinical evaluations are more discriminating than the test itself, they question why the test is being performed. One could make the same complaint to the authors: why do a PSG to diagnose OSA if 99% will be diagnosed with it using Stanford scoring criteria? Tapia et al. studied the influence of maturation on apnea/hypopnea duration and respiratory rates in 68 healthy non-obese 8- to 18-year-old controls (mean age 13 ± 3 years) using the AASM pediatric scoring rules, 32 of whom were adolescents (Tanner stages 3-5).[27] They re-scored the PSGs in the adolescents using AASM adult respiratory scoring criteria. These were normal controls, so they had very few respiratory events using any scoring criteria (median AHI 0/h using any criteria). The investigators found: (1) the majority of the few respiratory events present lasted ≥ 10 sec; and (2) the mean respiratory rate for the entire group was 17 ± 2 breaths per minute during NREM and REM sleep. A total of 36 hypopneas would have been scored in 15 of 32 adolescents using pediatric rules, only 2 using adult criteria (most often because these did not cause a ≥ 4% desaturation, only occasionally because they lasted < 10 sec). The authors argue that 2 missed breaths (regardless of duration) would be a more accurate criterion for scoring respiratory events at any age.

Differences in Sleep Architecture When Scoring Sleep in Adults Using AASM and R & K Rules

Moser et al. had 7 sleep experts from 3 European sleep laboratories re-score 72 adult PSGs using AASM sleep scoring criteria[28] and compared their results to R & K scorings (based upon consensus) done years earlier as part of the European SIESTA database.[29,30] When scored using AASM sleep scoring rules, they found statistically different (but clinically *insignificant)* increases in mean Wake after Sleep Onset (WASO, +4 min), NREM 1 (+11 min, +3% TST) and NREM 3 (+9 min, +2% TST) and decreased NREM 2 sleep (-21 min, -5% TST).

The authors further analyzed *why* particular epochs were re-scored using the AASM rules. Two-thirds rescored as Wake had been scored as Movement Time; the remaining third because sleep onset was defined as the first epoch of any stage of sleep

(vs. the first 3 epochs of stage 1 or any other stage of sleep in the SIESTA database). The AASM rules require slow wave activity of NREM 3 to be preferentially scored in the frontal EEG derivations; R & K restricted to the central regions. The modest increase in percent NREM 3 occurred because some had slow wave activity that was considerably greater than 75 μV in the frontal EEG derivations, but slightly less than 75 μV in the central.

The greatest change in sleep architecture following the AASM sleep scoring rules was a mean decrease in NREM 2 sleep (5%). This most often occurred because AASM rule 5.C.b was consistently applied, which specifies stop scoring successive epochs as NREM 2 sleep after an arousal and not score subsequent epochs as NREM 2 until a sleep spindle or K-complex (without an arousal) occurs. The 5.C.b. rule decreased NREM 2 time, increased NREM 1 time, and provided a subtle measure of sleep fragmentation by signaling a sleep stage shift. **Figure 1** shows 60 sec of a PSG illustrating application of this rule.

The 5.C.b. rule has resulted in a renewed interest in the number of sleep stage shifts which occur in a PSG. Laffan et al. retrospectively analyzed whether sleep stage shifts (which they call sleep stage transitions) were predictive of self-reported restless and light sleep in 5,684 participants in the Sleep Heart Sleep Study database.[31] Multivariable regression models showed that a high overall sleep stage transition rate was associated with restless and light sleep independent of TST, WASO, arousal index, and percentages of sleep stages. The authors found: (1) more than 10 sleep stage shifts per hour of sleep were associated with a 1.4 odds ratio of complaints of restless or light sleep; and (2) complaints of restless/light sleep increased linearly with the stage shift rate, plateauing when > 12 shifts/h.

Ruehland et al. had 3 experienced scorers from different Australian sleep laboratories score sleep stages and arousals in 30 adults with suspected OSA (mean age 51 years, BMI 34).[31] The PSGs were de-identified and presented to the scorers in random order to eliminate any order effect. When they scored PSGs using AASM rules, they found the mean time in NREM 1 decreased (mean 80 vs. 71 min) and time in NREM 3 increased (mean 64 vs. 53 min). They scored more NREM 3 (at the expense of NREM 2) because of slow wave activity being preferentially higher over the frontal regions. They scored more NREM 2 (at the expense of NREM 1) because K-complexes were more often recognized and maximal over the frontal areas.[32-34]

Differences in Sleep Architecture When Scoring Sleep in Children Using AASM and R & K Modified for Children Criteria

Three recently published studies compare sleep stage scoring in pediatric subjects using the age-specific AASM scoring rules and consensus-based R & K rules modified for children.[35-37] Novelli et al. had 3 experienced pediatric sleep experts score PSGs from 45 healthy normal children (mean age 8.5 ± 3 years)[35] using consensus-based R & K criteria adapted for children.[38] Two months later after a training lesson, they rescored them using the AASM pediatric sleep scoring rules.[23] They found they scored more NREM 1, more stage shifts, and less NREM 2 and REM sleep using the AASM rules (**Table 1**). Further analysis showed the increase in NREM 1 and stage shifts occurred when they applied rule 5.C.b. The percent or time in NREM 3 did not

Figure 1—NREM 1 sleep time and sleep stage shifts increase when scoring sleep studies using AASM sleep scoring rule 5.C.b.
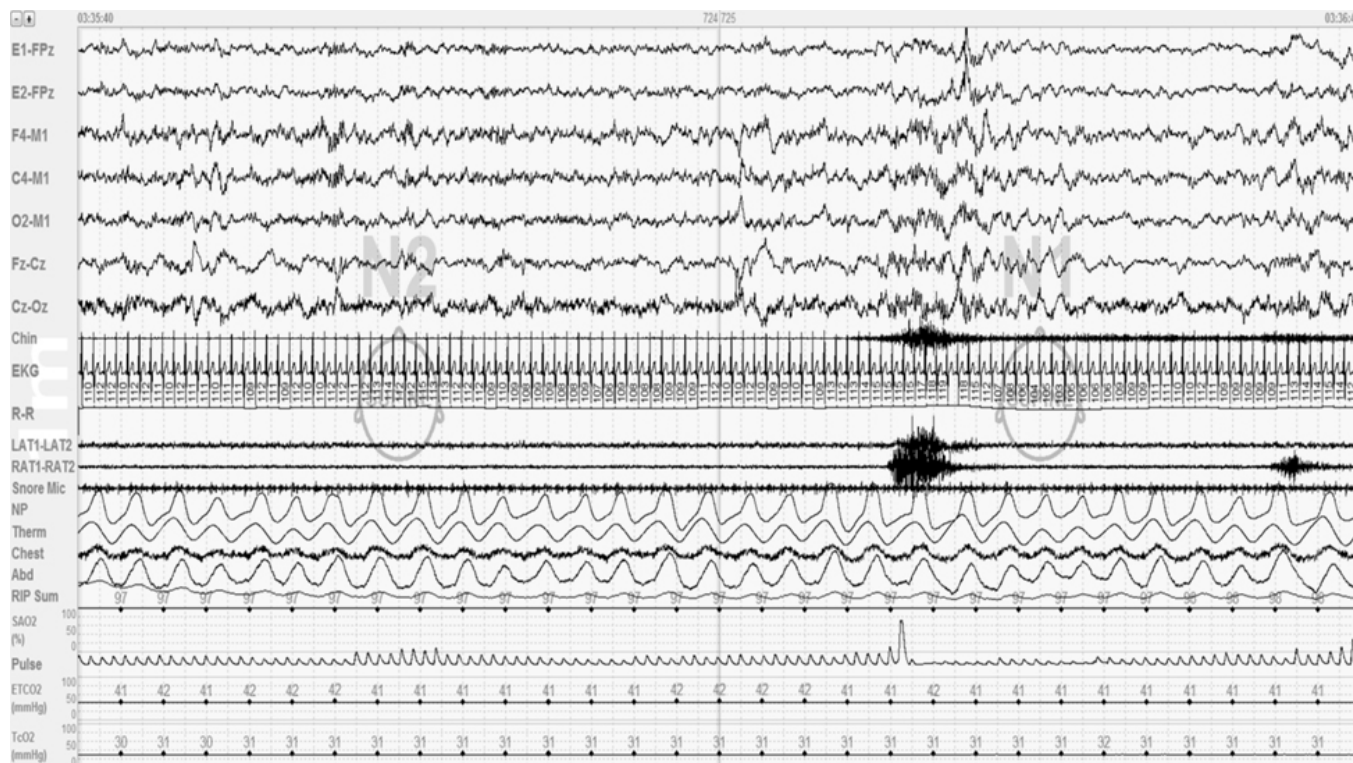

Figure showing a 60-sec epoch from a polysomnogram done on one of my patients. In the first 30-sec epoch, he is in NREM 2 sleep, then has an arousal (arrow). Because no sleep spindles or K-complexes recur following the arousal, the second epoch is scored as NREM 1 because of AASM scoring rule 5.C.b (p. 26 of the AASM Manual). Rule 5.C.b. has proved useful, increasing the amount and percent time of NREM 1 sleep and sleep stage shifts scored, providing a clinically useful marker of sleep fragmentation. This figure is from my laboratory, no permission needed.

Table 1—Significant differences in sleep study statistics when scoring pediatric polysomnograms using AASM and R & K rules adapted for children from study by Novelli et al.[35]

Sleep Stage Statistic Difference	AASM (Mean, % TST)	R & K Adapted for Children (Mean, % TST)
NREM 1 sleep	50 min (+10% TST)	18 min (4% TST)
NREM 2 sleep	236 min (46% TST)	263 min (51% TST)
REM sleep	110 min (21% TST)	120 min (23% TST)
Stage shifts per hour	9 ± 2/h	7 ± 2/h

increase because slow wave activity in children was typically > 150 to 400 µV over either the frontal and/or central regions, far greater voltages than the > 75 µV needed to score it.[39] The authors concluded the increased NREM 1 time and resultant stage shifts using AASM sleep scoring rules provided useful polysomnographic measures for fragmentation of sleep in children, not seen when scoring pediatric PSGs using R & K.

Miano et al. prospectively studied whether the AASM sleep scoring rules better differentiated sleep fragmentation in children with OSA or primary snoring from controls.[36] They scored PSGs in 75 prepubertal children (39 with OSA defined as an OAHI > 1: 26 primary snorers and 10 age-matched normal controls) using AASM and modified R & K.[36] They found the AASM rules better differentiated children with OSA from primary snorers or healthy age-matched controls (**Table 2**). Using the AASM rules, children with OSA had more NREM 1 and stage shifts per hour of sleep than primary snorers or controls

(again because of rule 5.C.b). The more severe the OSA, the higher the percent time spent in NREM 1 using the AASM (but not the R &K) criteria. Primary snorers could also be distinguished from controls by a significantly higher percent of NREM 1 sleep time using AASM (but not R & K) criteria.

Two studies have been published which provide normative first-night PSG data using the AASM rules for scoring sleep in infants and children.[37,40] The prospective longitudinal study by Sankupellay et al. confirmed that stages NREM 2 and 3 sleep could be scored in infants 3 months or older.[37] **Figure 2** shows sleep spindles best seen in the Fz-Cz channel recorded in a 3-month-old infant.

Interscorer Reliability Improved Using AASM Sleep Scoring Criteria

The AASM Visual Scoring of Adults review paper[41] cited 7 studies as evidence regarding IRR of humans to visually score

Table 2—Percent of agreement scoring polysomnograms using adult AASM and R & K sleep stage scoring rules according to study by Danker-Hopfe et al.[83]

Percent agreement of PSGs scored with different rules	Percent agreement between individual scorers using AASM	Percent agreement between consensus scored R & K
Excellent agreement (κ > 0.80)	42%	23%
Substantial agreement (κ = 0.61-0.79)	42%	66%
Moderate agreement (κ = 0.41-0.59)	14%	11%

Figure 2—Sleep spindles in a 3-month-old infant
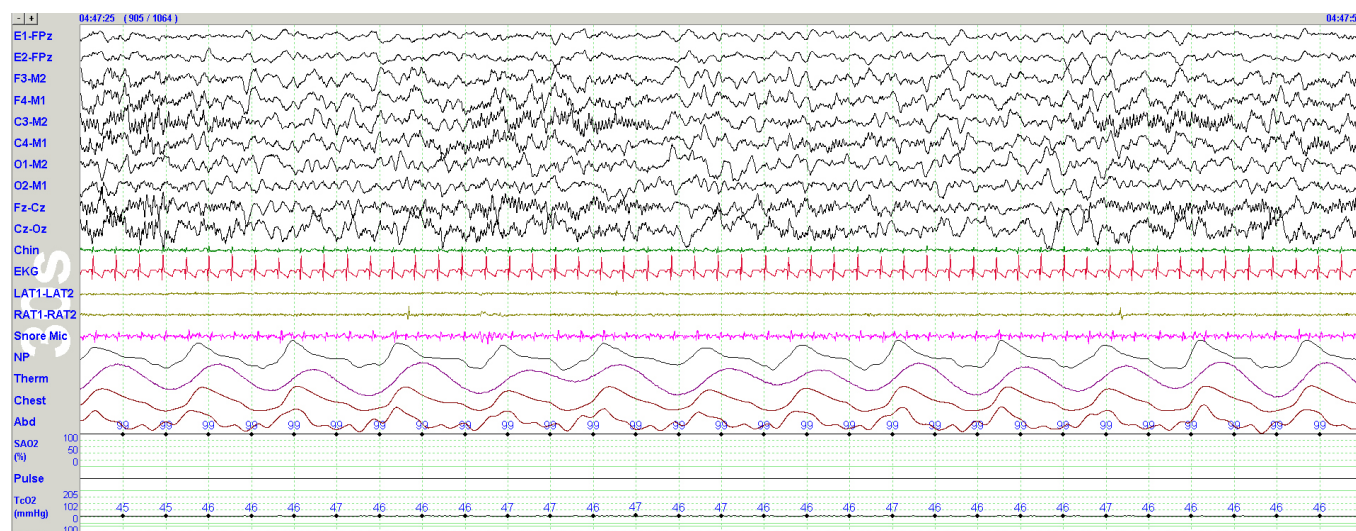
Note well-developed sleep spindles in this 30-sec epoch from a PSG recorded on a 3-month-old infant. Sleep spindles between 3-4 months of age often last 5-8 sec and are maximal over Cz. Sensitivity 500 μV peak-to-peak, LFF 0.3 Hz, HFF 35 Hz.

Table 3—AASM Pediatric sleep stage scoring rules better identified abnormalities in sleep architecture in children with obstructive sleep apnea shown in a study by Miano et al. [36]

	1) Controls (n = 10)	2) Primary Snorers (n = 26)	3) OSA (n = 39)	Statistical Significance (p <)
NREM 1 (% of TST)	8 ± 5 %	11 ± 4 %	14 ± 5%	1 vs. 2: 0.01 1 vs. 3: 0.005 2 vs. 3: 0.01
NREM 3 (% of TST)	37 ± 4%	39 ± 7%	32 ± 9%	1 vs. 2: NS 1 vs. 3: 0.05 2 vs. 3: 0.005
Stage stage shift (n/h)	12 ± 3/h	12 ± 4/h	16 ± 8/h	1 vs. 2: NS 1 vs. 3: 0.05 2 vs. 3: 0.005

sleep in a PSG using R & K rules,[15,42-47] including 2 graded level I,[15,44] and 2 level II evidence.[46,47] When PSGs were scored using R & K rules, interscorer agreement was highest for REM (78%-94%) and stage 2 (79%-90%) sleep; 68%-89% for Wake; 23%-74% stage 1, 44%-60% stage 3; 45%-80% stage 4; and 69% for slow wave sleep (stages 3 and 4 combined).[15,44,48-51] NREM 2 and NREM 3 were most likely to be interchanged, followed by stage NREM 1 and NREM 2.[52]
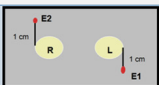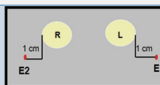
Two studies have been published which evaluate whether IRR improved when sleep was scored using AASM sleep scoring criteria.[31,53] Danker-Hopfe et al. took those same 72 SIESTA database PSGs that had been scored by consensus[28] and com-pared them with "first scorings" by individuals using AASM rules.[53] Using arbitrary agreement benchmarks according to Landis and Koch, κ agreement was perfect (κ > 0.80) in 42% of the Manual scoring vs. 23% scored using R &K (**Table 3**). Using Cohen and Koch κ coefficients, the overall IRR was statistically higher using AASM (82% agreement, κ = 0.75, p = 0.02) vs. 80.6% (κ = 0.72) following R & K. Percent scorer agreements increased using AASM for all sleep stages save NREM 2 (again because of rule 5.C.b). Discrepancies in scoring were found in 18% of all epochs using Manual rules compared with 19% using R & K; 60% of these were between slow wave sleep/NREM 3 and stage 2/NREM 2. Levels of agreement decreased with in-

Figure 3—Cancellation effects of the bipolar linkage Fz-Cz on slow wave activity of NREM 3 sleep
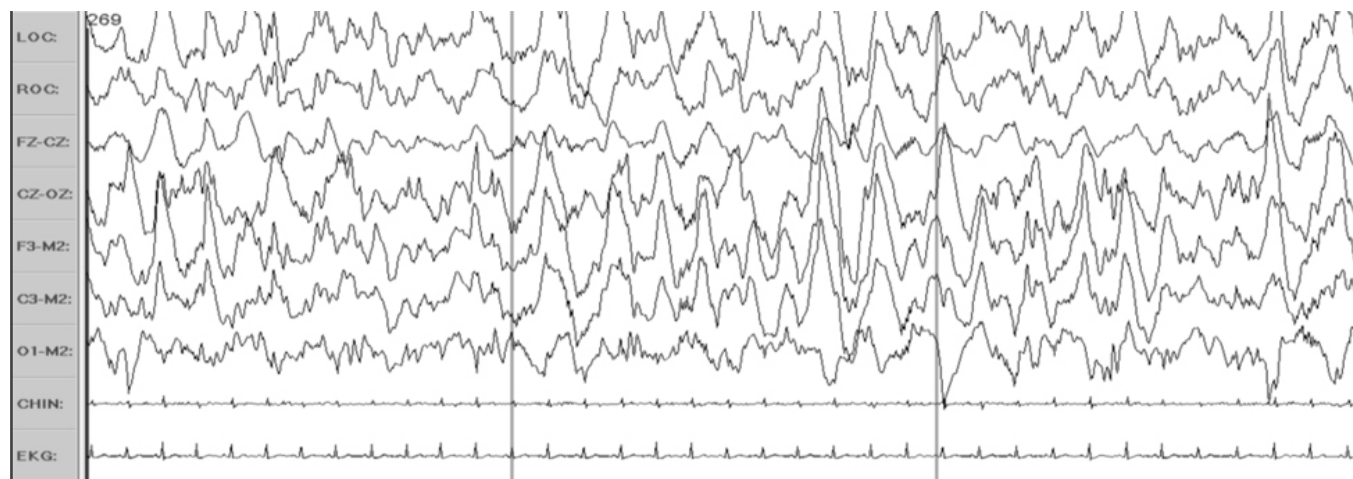
Note how much lower the amplitude is in the Fz-Cz and Cz-Oz channels compared to F3-M2, C3-M2, and O1-M2 due to cancellation effects of bipolar linkages over frontocentral regions where slow wave activity is often synchronous and of each amplitude.

Figure 4—Advantages and disadvantages of recommended and alternative EOG montages

creasing age of the patient, and this was more pronounced in the R & K scorings (p = -0.008). The authors compared interscorer agreement epoch-by-epoch between 2 epochs; agreement was greatest for REM, followed by Wake, NREM 3, NREM 2, and NREM 1 sleep. In summary, interscorer reliability in this study was greater using AASM rules.

However, the study by Ruehland et al. found no statistically significant differences in interscorer or intrascorer reliability for sleep or arousals using AASM and R & K sleep scoring criteria.[54] They reported a trend for better IRR scoring sleep and arousals using the AASM rules and questioned small sample size and/or that their study population were patients (rather than normal controls), led to the lack of significant differences between criteria.

Technical Considerations Recommended by the AASM Manual

The AASM Scoring Manual allows for 2 different EEG and EOG montages. Why two of each? The recommended EEG montage is similar to R & K, except for adding electrodes over the frontal and occipital regions. The recommended EEG montage references the frontal, central, and occipital electrodes to

contralateral mastoid—referential derivations are the best way to measure amplitude in EEG. The occipital EEG derivations were added because the dominant posterior alpha rhythm used to identify stage W is usually maximal over this region. Frontal derivations were added because K-complexes and NREM 3 slow wave activity are often of highest amplitude there. Sleep spindles, vertex waves, and sawtooth waves are usually best seen over the central EEG derivations.

The alternative EEG and EOG montages have been used for years at the Mayo Clinic. Advantages of the alternative EEG montage cited by its proponents were: (1) the PSG signatures of sleep are maximal over Fz, Cz, and Oz; (2) linking biologically active electrodes (Fz to Cz, Cz to Oz) often results in "cleaner" EEG signals; and (3) Fz-Cz often useful for identifying electrographic seizure activity because of less muscle artifact over midline central regions. A single study by van Sweden et al. found overall intra- and interscorer agreement were similar when sleep stages in a PSG were scored using Fz-Cz/Cz-Pz or C4-A1.[55] They further reported a trend for more reliable scoring of NREM 3 using the alternative EEG placement.

However, a disadvantage when linking Fz to Cz in the alternative EEG montage is that it can produce cancellation effects when NREM 3 slow wave activity or K-complexes are of similar amplitude over the frontal and central regions. **Figure 3** shows an example of cancellation effects using Fz-Cz to score slow activity of NREM 3. Soon after the Manual was published, queries appeared in the Frequently Asked Questions for the Manual website, asking how to score NREM 3 using the alternative EEG montage. The Manual Steering Committee recommended E1-FPz when using the alternative EOG montage (which can work) or C4-M1 if using recommended EOG derivation (which does not solve the problem of using a frontal derivation). Fz-M1 might also suffice. **Figure 4** summarizes the advantages and disadvantages of recommended and alternative EOG montages. **Figure 5** illustrates how the alternative EOG montage is much more sensitive than the recommended for detecting eye movements

Figure 5—The alternative EOG montage better identifies eye movements

Note how the alternative EOG montage better demonstrates the direction of eye movements and subtle eye blinks compared with the recommended EOG montage (Figure provided by Michael Silber, M.D.)

Other Complaints, Requests for Clarification, and Suggested Revisions

The Italian Association of Sleep Medicine voiced concern that only 3 of the 29 rules for visual scoring of sleep had sufficient evidence to warrant a recommendation as a standard, that the vast majority were based on consensus.[56] They questioned discarding PSG data collected for years without first performing multicenter validation studies to confirm the rules and recommendations, especially "given the low levels of evidence offered as reasoning."

Schulz criticized the AASM sleep scoring rules because they reduce the "multilayer and time-varying physiologic processes of sleep" to a "rule-based set of a few sleep stages" and ignore "shifting levels of vigilance and arousal across a night of sleep".[57] The AASM Visual Scoring Task Forces reviewed and debated the evidence for adopting alternative methods such as adaptive segmentation, cluster quasi-stationary segments of varying duration, or a wavelet transform in which the signal is localized in the time-frequency plane. However, compelling evidence remains to be published then or now that alternative methods that are usually more time-consuming improve clinical outcome or diagnostic acumen than fixed duration 30-sec epoch-by-epoch scoring.[58]

Miano et al. questioned why we fail to provide rules for scoring cortical arousals from NREM 3.[36] **Figure 6** shows an arousal from NREM 3 sleep in a 13-year-old child. Pediatric electroencephalographers would identify this as a hypnapompic EEG arousal and the EEG background shift into the delta EEG frequency range as

NREM 1. The pediatric scoring task force debated at length as to whether EEG shifts into delta frequencies should be scored as an arousal. They voted no because the IRR for scoring delta EEG shifts from NREM sleep was suboptimal and because they were uncertain how "arousing" a delta EEG arousal was.

Scholle et al. propose we revise the rules for scoring arousals in children to permit a second arousal to be scored after only 5 (not 10) seconds of stable sleep because respiratory events in young children often last only 5.5 to 8.5 seconds.[40] Last but not least, several pediatric sleep specialists decry lumping stages 3 and 4 into NREM 3 sleep[35,36,40] because it dilutes it as a robust marker of homeostatic drive, sleep debt, pathology, and sleep development.[59-64]

No studies have been published evaluating or contesting the validity or reliability of AASM rules for scoring periodic limb movements of sleep, cardiac events, or REM sleep without atonia in a PSG. Researchers are studying which muscles are best to record REM sleep without atonia in a PSG, the cut-off points for it, and whether changes in heart rate and periodicity of PLMS can differentiate limb movements in patients with restless legs from those seen in narcolepsy with cataplexy, REM sleep behavior disorder, and asymptomatic older adults.[65-68]

In Closing

Four years have passed since the AASM Scoring Manual was published with far less criticism than those who developed it feared.[69] Perhaps the paucity of printed criticism reflects the

Figure 6—An EEG arousal from NREM 3 sleep in a 13-year-old
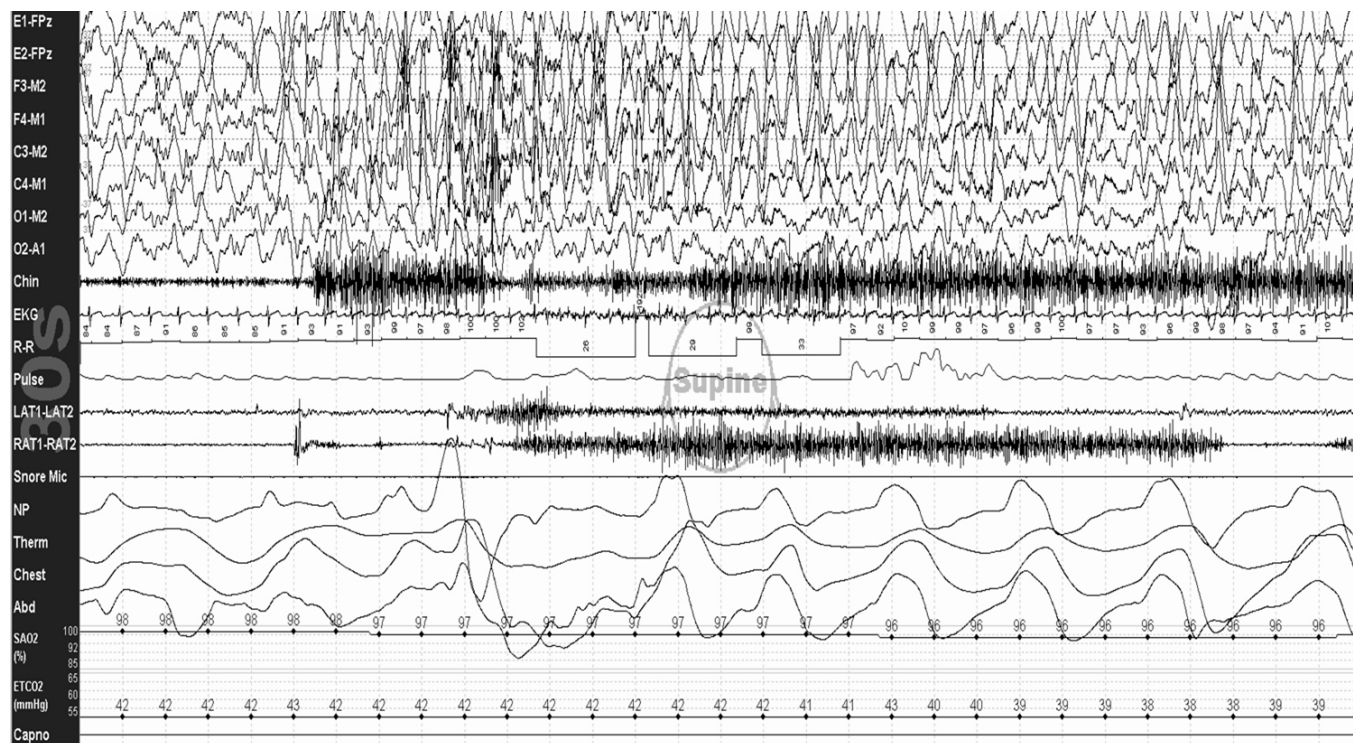
An arousal from NREM 3 sleep in a 13 year old. Pediatric electroencephalographers would score this EEG delta shift as NREM 1 sleep.

realization that we now have a comprehensive set of rules and recommendations which we can with further research challenge, discard, expand upon, and revise. Sufficient evidence favors a single rule for scoring hypopneas in adults. The improvement in interscorer reliability using the AASM rules is encouraging, but more studies are needed to confirm this.

Critics of the Manual disparage its explicit simplicity. However, if at least 85% to 90% of PSGs are done for suspected obstructive sleep apnea and the AASM rules identify it in the majority of these, why score more? The rules will miss some whose symptoms are truly related to their sleep disorder(s), but given increasing pleas for medical economic restraint, full-court press PSG techniques should perhaps be reserved for those whose first study is thought to be false-negative or non-diagnostic. Studies remain to be done and published which confirm that scoring subcortical arousals, cyclic alternating pattern,[70-74] respiratory cycle-related EEG changes to identify respiratory events,[75-77] abnormal nocturnal regulation of cardiac autonomic function,[78] computerized analysis of submentalis muscle activity to diagnose REM sleep behavior disorder,[79] and assessment of periodicity and circadian distribution of periodic limb movements[65,66,68,80-82] are warranted, valid, and alter outcomes.

REFERENCES

1. Iber C, American Academy of Sleep Medicine. *The AASM manual for the scoring of sleep and associated events : rules, terminology and technical specifications.* Westchester, IL: American Academy of Sleep Medicine; 2007.
2. Iber C, Ancoli-Israel S, Chambers M, Quan SF. The new sleep scoring manual-the evidence behind the rules. *J Clin Sleep Med* 2007;3:107.
3. Iber C. Development of a new manual for characterizing sleep. *Sleep* 2004;27:190-2.
4. Fitch K. *The Rand/UCLA appropriateness method user's manual.* Santa Monica: Rand; 2001.
5. Redline S, Budhiraja R, Kapur V, et al. The scoring of respiratory events in sleep: reliability and validity. *J Clin Sleep Med* 2007;3:169-200.
6. Young T, Palta M, Dempsey J, Skatrud J, Weber S, Badr S. The occurrence of sleep-disordered breathing among middle-aged adults. *N Engl J Med* 1993;328:1230-5.
7. Young T, Peppard P, Palta M, et al. Population-based study of sleep-disordered breathing as a risk factor for hypertension. *Arch Intern Med* 1997;157:1746-52.
8. Kim HC, Young T, Matthews CG, Weber SM, Woodward AR, Palta M. Sleep-disordered breathing and neuropsychological deficits. A population-based study. *Am J Respir Crit Care Med* 1997;156:1813-9.
9. Young T, Blustein J, Finn L, Palta M. Sleep-disordered breathing and motor vehicle accidents in a population-based sample of employed adults. *Sleep* 1997;20:608-13.
10. Nieto FJ, Young TB, Lind BK, et al. Association of sleep-disordered breathing, sleep apnea, and hypertension in a large community-based study. Sleep Heart Health Study. *JAMA* 2000;283:1829-36.
11. Shahar E, Whitney CW, Redline S, et al. Sleep-disordered breathing and cardiovascular disease: cross-sectional results of the Sleep Heart Health Study. *Am J Respir Crit Care Med* 2001;163:19-25.
12. Gottlieb DJ, Yao Q, Redline S, Ali T, Mahowald MW. Does snoring predict sleepiness independently of apnea and hypopnea frequency? *Am J Respir Crit Care Med* 2000;162(4 Pt 1):1512-7.
13. Hirshkowitz M, Sharafkhaneh A. On measurement consistency and other hobgoblins. *Sleep* 2004;27:847-8.
14. Ayappa I, Norman RG, Krieger AC, Rosen A, O'Malley RL, Rapoport DM. Non-Invasive detection of respiratory effort-related arousals (RERAs) by a nasal cannula/pressure transducer system. *Sleep* 2000;23:763-71.
15. Whitney CW, Gottlieb DJ, Redline S, et al. Reliability of scoring respiratory disturbance indices and sleep staging. *Sleep* 1998;21:749-57.
16. American Academy of Sleep Medicine. Scoring Manual Frequently Asked Questions. 2007. Accessed January 16, 2012, 2012. http://www.aasmnet.org/scoringmanualfaq.aspx

17. Ruehland WR, Rochford PD, O'Donoghue FJ, Pierce RJ, Singh P, Thornton AT. The new AASM criteria for scoring hypopneas: impact on the apnea hypopnea index. *Sleep.* Feb 1 2009;32:150-7.

18. Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research. The Report of an American Academy of Sleep Medicine Task Force. *Sleep* 1999;22:667-89.

19. Meoli AL, Casey KR, Clark RW, et al. Hypopnea in sleep-disordered breathing in adults. *Sleep* 2001;24:469-70.

20. Guilleminault C, Hagen CC, Huynh NT. Comparison of hypopnea definitions in lean patients with known obstructive sleep apnea hypopnea syndrome (OSAHS). *Sleep Breath* 2009;13:341-7.

21. Norman MB, Middleton S, Sullivan CE. The use of epochs to stage sleep results in incorrect computer-generated AHI values. *Sleep Breath* 2010.

22. Praetorius HM, Bodenstein G, Creutzfeldt OD. Adaptive segmentation of EEG records: a new approach to automatic EEG analysis. *Electroencephalogr Clin Neurophysiol* 1977;42:84-94.

23. Iber C, Ancoli-Israel S, Chessonn A, Quan SF. *The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications.* 1st ed. Westchester, Illinois: American Academy of Sleep Medicine; 2007.

24. Accardo JA, Shults J, Leonard MB, Traylor J, Marcus CL. Differences in overnight polysomnography scores using the adult and pediatric criteria for respiratory events in adolescents. *Sleep* 2010;33:1333-9.

25. Lin CH, Guilleminault C. Current hypopnea scoring criteria underscore pediatric sleep disordered breathing. *Sleep Med* 2011;12:720-9.

26. American Academy of Sleep Medicine. *The international classification of sleep disorders : diagnostic and coding manual.* 2nd ed. Westchester, Ill.: American Academy of Sleep Medicine; 2005.

27. Tapia IE, Karamessinis L, Bandla P, et al. Polysomnographic values in children undergoing puberty: pediatric vs. adult respiratory rules in adolescents. *Sleep* 2008;31:1737-44.

28. Moser D, Anderer P, Gruber G, et al. Sleep classification according to AASM and Rechtschaffen & Kales: effects on sleep scoring parameters. *Sleep* 2009;32:139-49.

29. Rechtschaffen A, Kales A. *A manual of standardized terminology, techniques, and scoring system for sleep stages of human subjects.* Bethesda, MD: National Institute of Neurological Disease and Blindness; 1968.

30. Klosch G, Kemp B, Penzel T, et al. The SIESTA project polygraphic and clinical database. *IEEE Eng Med Biol Mag* 2001;20:51-7.

31. Laffan A, Caffo B, Swihart BJ, Punjabi NM. Utility of sleep stage transitions in assessing sleep continuity. *Sleep* 2010;33:1681-6.

32. McCormick L, Nielsen T, Nicolas A, Ptito M, Montplaisir J. Topographical distribution of spindles and K-complexes in normal subjects. *Sleep* 1997;20:939-41.

33. 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.

34. Werth E, Achermann P, Borbely AA. Fronto-occipital EEG power gradients in human sleep. *J Sleep Res* 1997;6:102-12.

35. Novelli L, Ferri R, Bruni O. Sleep classification according to AASM and Rechtschaffen and Kales: effects on sleep scoring parameters of children and adolescents. *J Sleep Res* 2010;19(1 Pt 2):238-47.

36. Miano S, Paolino MC, Castaldo R, Villa MP. Visual scoring of sleep: A comparison between the Rechtschaffen and Kales criteria and the American Academy of Sleep Medicine criteria in a pediatric population with obstructive sleep apnea syndrome. *Clin Neurophysiol* 2010;121:39-42.

37. Sankupellay M, Wilson S, Heussler HS, Parsley C, Yuill M, Dakin C. Characteristics of sleep EEG power spectra in healthy infants in the first two years of life. *Clin Neurophysiol* 2011;122:236-43.

38. Scholle S, Schafer T. Atlas of states of sleep and wakefulness in infants and children. *Somnologie* 1999;3:163-241.

39. Grigg-Damberger M, Gozal D, Marcus CL, et al. The visual scoring of sleep and arousal in infants and children. *J Clin Sleep Med* 2007;3:201-40.

40. Scholle S, Beyer U, Bernhard M, et al. Normative values of polysomnographic parameters in childhood and adolescence: quantitative sleep parameters. *Sleep Med* 2011;12:542-9.

41. Silber MH, Ancoli-Israel S, Bonnet MH. The visual scoring of sleep in adults. *J Clin Sleep Med* 2007;3:121-31.

42. Martin W, LC Johnson, SS Viglione, et al. Pattern recognition of EEG–EOG as a technique for all-night sleep stage scoring. *Electroencephalogr Clin Neurophysiol* 1972;32:417-27.

43. Kim Y, Kurachi M, Horita M, Matsuura K, Kamikawa Y. Agreement of visual scoring of sleep stages among many laboratories in Japan: effect of a supplementary definition of slow wave on scoring of slow wave sleep. *Jpn J Psychiatry Neurol* 1993;47:91-7.

44. Danker-Hopke 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.

45. Monroe L. Inter-rater reliability and the role of experience in scoring EEG sleep. *Psychophysiology* 1967;5:376-84.

46. Norman RG, Pal I, Stewart C, Walsleben JA, Rapoport DM. Interobserver agreement among sleep scorers from different centers in a large dataset. *Sleep* 2000;23:901-8.

47. Sangal RB, Semery JP, Belisle CL. Computerized scoring of abnormal human sleep: a validation. *Clin Electroencephalogr* 1997;28:64-7.

48. Feinberg I, Thode HC, Jr., Chugani HT, March JD. Gamma distribution model describes maturational curves for delta wave amplitude, cortical metabolic rate and synaptic density. *J Theor Biol* 1990;142:149-61.

49. Sangal RB, Semery JP, Belisle CL. Computerized scoring of abnormal human sleep: a validation. *Clin Electroencephalogr* 1997;28:64-7.

50. Kim Y, Kurachi M, Horita M, Matsuura K, Kamikawa Y. Agreement of visual scoring of sleep stages among many laboratories in Japan: effect of a supplementary definition of slow wave on scoring of slow wave sleep. *Jpn J Psychiatry Neurol* 1993;47:91-7.

51. Martin WB, Johnson LC, Viglione SS, Naitoh P, Joseph RD, Moses JD. Pattern recognition of EEG-EOG as a technique for all-night sleep stage scoring. *Electroencephalogr Clin Neurophysiol* 1972;32:417-27.

52. Kirkness JP, Schwartz AR, Schneider H, et al. Contribution of male sex, age, and obesity to mechanical instability of the upper airway during sleep. *J Appl Physiol* 2008;104:1618-24.

53. Danker-Hopfe H, Anderer P, Zeitlhofer J, et al. Interrater reliability for sleep scoring according to the Rechtschaffen & Kales and the new AASM standard. *J Sleep Res* 2009;18:74-84.

54. Ruehland WR, O'Donoghue FJ, Pierce RJ, et al. The 2007 AASM recommendations for EEG electrode placement in polysomnography: impact on sleep and cortical arousal scoring. *Sleep* 2011;34:73-81.

55. van Sweden B, Kemp B, Kamphuisen HA, Van der Velde EA. Alternative electrode placement in (automatic) sleep scoring (Fpz-Cz/Pz-Oz versus C4-A1). *Sleep* 1990;13:279-83.

56. Parrino L, Ferri R, Zucconi M, Fanfulla F. Commentary from the Italian Association of Sleep Medicine on the AASM manual for the scoring of sleep and associated events: For debate and discussion. *Sleep Med* 2009;10:799-808.

57. Schulz H. Rethinking sleep analysis. *J Clin Sleep Med* 2008;4:99-103.

58. Penzel T, Hirshkowitz M, Harsh J, et al. Digital analysis and technical specifications. *J Clin Sleep Med* 2007;3:109-20.

59. Buchmann A, Ringli M, Kurth S, et al. EEG sleep slow-wave activity as a mirror of cortical maturation. *Cereb Cortex* 2011;21:607-15.

60. Greene RW, Frank MG. Slow wave activity during sleep: functional and therapeutic implications. *Neuroscientist* 2010;16:618-33.

61. Le Bon O, Chabanski S, Dramaix M, Staner L, Pelc I, Linkowski P. Inverse association between slow wave activity per cycle and the number of ultradian sleep cycles per night in healthy humans. *Clin Neurophysiol* 2005;116:1493-500.

62. Jenni OG, Achermann P, Carskadon MA. Homeostatic sleep regulation in adolescents. *Sleep* 2005;28:1446-54.

63. Ohayon MM, Carskadon MA, Guilleminault C, 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.

64. Jenni OG, Carskadon MA. Spectral analysis of the sleep electroencephalogram during adolescence. *Sleep* 2004;27:774-83.

65. Ferri R, Zucconi M. Heart rate and spectral EEG changes accompanying periodic and isolated leg movements during sleep. *Sleep* 2008;31:16-7; discussion 18-9.

66. Ferri R, Manconi M, Lanuzza B, et al. Age-related changes in periodic leg movements during sleep in patients with restless legs syndrome. *Sleep Med* 2008;9:790-8.

67. Manconi M, Ferri R, Zucconi M, Fantini ML, Plazzi G, Ferini-Strambi L. Time structure analysis of leg movements during sleep in REM sleep behavior disorder. *Sleep* 2007;30:1779-85.

68. Ferri R, Zucconi M, Manconi M, et al. Different periodicity and time structure of leg movements during sleep in narcolepsy/cataplexy and restless legs syndrome. *Sleep* 2006;29:1587-94.

69. Grigg-Damberger MM. The AASM scoring manual: a critical appraisal. *Curr Opin Pulm Med* 2009. September 4 (Epub ahead of print)

70. Parrino L, Ferri R, Bruni O, Terzano MG. Cyclic alternating pattern (CAP): The marker of sleep instability. *Sleep Med Rev* 2012;16:27-45

71. Roach GD, Navarro DC, Sargent C. The evidence that cyclic alternating pattern subtypes affect cognitive functioning is very weak. *Sleep Med* 2010;11:803; author reply 803-4.
72. Bruni O, Novelli L, Miano S, Parrino L, Terzano MG, Ferri R. Cyclic alternating pattern: A window into pediatric sleep. *Sleep Med* 2010;11:628-36.
73. Smerieri A, Parrino L, Agosti M, Ferri R, Terzano MG. Cyclic alternating pattern sequences and non-cyclic alternating pattern periods in human sleep. *Clin Neurophysiol* 2007;118:2305-13.
74. Thomas RJ. Cyclic alternating pattern in the electroencephalogram: what is its clinical utility? *Sleep* 2007;30:553-5.
75. Chervin RD, Malhotra RK, Burns JW. Respiratory cycle-related EEG changes during sleep reflect esophageal pressures. *Sleep* 2008;31:1713-20.
76. Chervin RD, Burns JW, Subotic NS, Roussi C, Thelen B, Ruzicka DL. Correlates of respiratory cycle-related EEG changes in children with sleep-disordered breathing. *Sleep* 2004;27:116-21.
77. Chervin RD, Burns JW, Subotic NS, Roussi C, Thelen B, Ruzicka DL. Method for detection of respiratory cycle-related EEG changes in sleep-disordered breathing. *Sleep* 2004;27:110-5.
78. Lombardi C, Parati G, Cortelli P, et al. Daytime sleepiness and neural cardiac modulation in sleep-related breathing disorders. *J Sleep Res* 2008;17:263-70.
79. Ferri R, Manconi M, Plazzi G, et al. A quantitative statistical analysis of the submentalis muscle EMG amplitude during sleep in normal controls and patients with REM sleep behavior disorder. *J Sleep Res* 2008;17:89-100.
80. Ferri R, Gschliesser V, Frauscher B, Poewe W, Hogl B. Periodic leg movements during sleep and periodic limb movement disorder in patients presenting with unexplained insomnia. *Clin Neurophysiol* 2009;120:257-63.
81. Ferri R, Zucconi M, Manconi M, et al. Computer-assisted detection of nocturnal leg motor activity in patients with restless legs syndrome and periodic leg movements during sleep. *Sleep* 2005;28:998-1004.
82. Ferri R, Zucconi M, Manconi M, Plazzi G, Bruni O, Ferini-Strambi L. New approaches to the study of periodic leg movements during sleep in restless legs syndrome. *Sleep* 2006;29:759-69.
83. Danker-Hopfe H, Anderer P, Zeitlhofer J, et al. Interrater reliability for sleep scoring according to the Rechtschaffen & Kales and the new AASM standard. *J Sleep Res* 2009;18:74-84.

SUBMISSION & CORRESPONDENCE INFORMATION

DISCLOSURE STATEMENT

This was not an industry supported study. The author has indicated no financial conflicts of interest.