

The AASM Recommended and Acceptable EEG Montages are Comparable for the Staging of Sleep and Scoring of EEG Arousals

Brett Duce, BSc (Hons); Conchita Rego, MSc; Jasmina Milosavljevic, MBBS, MPH; Craig Hukins, MBBS

Sleep Disorders Centre, Department of Respiratory & Sleep Medicine, Princess Alexandra Hospital, Woolloongabba, Queensland, Australia

SCIENTIFIC INVESTIGATIONS

Study Objective: To examine the measurement differences in sleep and EEG arousal statistics between the American Academy of Sleep Medicine (AASM) recommended EEG montage (F4-M1, C4-M1, O2-M1) and acceptable EEG montage (Fz-Cz, C4-M1, Oz-Cz).

Design: Prospective, blinded, randomized comparison.

Setting: Australian clinical sleep laboratory in a tertiary hospital.

Patients or Participants: 50 consecutive patients undertaking polysomnography (PSG) for the clinical suspicion of sleep disordered breathing.

Interventions: N/A

Measurements and Results: Patient EEGs were recorded using both the AASM recommended and acceptable EEG montages during the PSG. Two scorers were used to examine the difference in PSG statistics using the two EEG montages. The scorers analyzed the 50 studies using the two EEG montages. Ten of the studies were scored twice for each

montage by each scorer to calculate intra-scorer and inter-scorer agreement. No statistically significant differences were observed between the PSG statistics of the recommended and acceptable EEG montages. The recommended EEG montage had greater inter-scorer agreement but no difference in intra-scorer agreement.

Conclusions: This study demonstrates that the two EEG montages endorsed by the AASM Manual produce similar sleep and EEG arousal statistics.

Keywords: EEG, recommended montage, acceptable montage, sleep staging, EEG arousal scoring, polysomnography, sleep apnea, inter-scorer agreement, intra-scorer agreement, kappa

Citation: Duce B, Rego C, Milosavljevic J, Hukins C. The AASM recommended and acceptable EEG montages are comparable for the staging of sleep and scoring of EEG arousals. *J Clin Sleep Med* 2014;10(7):803-809.

A standard of clinical polysomnography (PSG) is the characterization of sleep and sleep fragmentation via the staging of sleep and scoring of EEG arousals.¹ For a number of decades the staging of sleep was conducted according to the Rechtschaffen and Kales (R&K) criteria.² This standard recommended a single central derivation, despite its apparent shortcomings.³ In 2007 the American Academy of Sleep Medicine (AASM) published *The AASM Manual for the Scoring of Sleep and Associated Events*.⁴ This was seen as a long-overdue update of the criteria, incorporating the evidence accumulated since the publication of R&K and recognizing the need to improve inter-scorer agreement.⁵

Among the changes in the AASM manual was the increase to three derivations in the EEG montage as well as the provision of two different EEG montages. The introduction of two extra derivations into the EEG montage was based on evidence that regional differences in EEG morphology exist and may be important in describing sleep.⁶⁻⁹ The two EEG montages allowed differ in the type of derivations that make up each montage. The recommended montage (M_{rec}) consists of all referential derivations, with negative input electrodes over the frontal, central, and occipital regions. The acceptable montage (M_{acc}), on the other hand, consists of two bipolar derivations (Fz-Cz and Cz-Oz) and a referential derivation (C4-M1). Despite the

BRIEF SUMMARY

Current Knowledge/Study Rationale: The American Academy of Sleep Medicine's Manual for the Scoring of Sleep and Associated Events allows for the use of either a recommended EEG montage or an acceptable EEG Montage in polysomnographic recordings. In this study we compared sleep and EEG arousals indices derived from the two montages to determine their equivalence.

Study Impact: There was very little difference between montages in sleep parameters such as total sleep time, EEG arousal index and sleep stage proportions. Greater inter-scorer agreement was observed with the recommended montage but likely reflects the scorers' greater familiarity with the montage.

“recommended” and “acceptable” nomenclature, the AASM has indicated that each montage is equally acceptable for use.¹⁰

While both montages are equally acceptable for use, little is known about the comparability of sleep statistics produced by each montage. The referential derivations which make up M_{rec} have the advantage of magnifying the waveform amplitudes close to the negative input electrodes, but they are also more prone to signal noise. The bipolar derivations which make up part of the M_{acc} are less likely to be affected by artifact but may potentially mask global EEG activity.¹¹ Furthermore, bipolar derivations tend to produce waveforms with smaller

amplitudes due to their shorter inter-electrode distance.¹² This is due to both the proximity of the electrode locations and the nature of differential amplifiers. Proximal electrodes will often receive similar potentials oriented in the same dipole,¹³ leading the differential amplifier to cancel these potentials out.

This EEG cancellation effect has the capacity to affect the scoring of slow wave sleep if the Fz-Cz in M_{acc} is used for measuring the amplitude of frontal activity. Kemp and colleagues have recently shown slow wave amplitudes in the Fz-Cz derivation to be approximately 60% to 65% of the slow wave amplitudes measured in the F4-M1 derivation.¹³ This could lead to significant changes in the proportion of stage 3 sleep (N3) scored. The AASM acknowledged this issue in item V5 of their online FAQ,¹⁰ and suggested using C4-M1 if using the recommended electrooculogram (EOG) placement and E1-Fpz if using the alternative EOG placement for the measurement of frontal activity. The correspondence between M_{rec} and M_{acc} for other sleep indices is, however, less certain. Other sleep indices, such as total sleep time and arousal index, are also important in the description of sleep and are often used to generate other important indices (e.g., apnea-hypopnea index [AHI]). Thus it is necessary to appreciate any differences between the two montages for a range of sleep indices.

The aim of this study was therefore to compare sleep and EEG arousal indices derived from the two AASM derivations (M_{rec} and M_{acc}) in a patient cohort examined for sleep disordered breathing.

METHODS

Patient Selection

Consenting patients undergoing diagnostic PSG for suspected sleep disordered breathing were recruited for the study. Patients were excluded if they had undergone diagnostic PSG previously or if a split-night treatment protocol (diagnostic to positive airway pressure therapy) was implemented. The institutional human research ethics committee approved this study.

Determination of Sample Size

While the relative proportion of sleep stages can be significant in clinical PSG, the correct determination of total sleep time (TST) was deemed to be paramount. The TST is often used in the calculation of other PSG statistics such as the AHI. To determine the sample size, we used TST as the primary outcome measure. Based on data from the previous 3 months of diagnostic PSGs, we calculated that a sample of 50 patients was necessary to detect a difference in TST of 15 minutes with a two-sided 5% significance level and a power of 80%.

Polysomnography

PSGs were recorded with the E-series acquisition system (Compumedics, Abbotsford, Australia). The recording montage comprised both M_{rec} and M_{acc} derivations (F4-M1, C4-M1, O2-M1, Fz-Cz, Cz-Oz), left and right EOG (recommended derivation: E1-M2, E2-M2), chin electromyogram (EMG, mental/submental positioning), modified lead II ECG, nasal pressure (DC amplified), oronasal thermocouple, body position, thoracic

and abdominal effort (inductive plethysmography), pulse oximetry (Nonin Xpod 3011), left and right leg movement (anterior tibialis EMG), and sound pressure (dBA meter: Tecpel 332).

Polysomnogram Scoring Protocol

Following acquisition, PSG's were de-identified and modified to display a staging montage with either M_{rec} or M_{acc} available for viewing. PSGs were presented to 2 scorers in randomized order. Randomization of PSGs was performed using the freely accessible Randomizer website.¹⁴ In order to calculate intra-scorer and inter-scorer agreement, the first 10 of the 50 PSGs were scored 4 times (twice using M_{rec} and twice using M_{acc}) by each scorer. The next 20 PSGs were scored twice by Scorer 1 (JM)—once each using M_{rec} and M_{acc} . The last 20 PSGs were scored by Scorer 2 (CR) in similar fashion. Overall, each scorer was presented with a total of 80 PSGs to score.

The staging montage consisted of 2 viewing panes. The upper pane displayed ECG, Chin EMG, left EOG, right EOG, EEG1 (frontal - F4-M1 or Fz-Cz), EEG2 (central - C4-M1) and EEG3 (occipital - O2-M1 or Oz-Cz) and was viewed across a 30-s epoch. A lower pane displayed the nasal pressure and left and right leg anterior tibialis EMG channels across a 5-min epoch. Sleep staging and EEG arousal scoring were performed according to the AASM manual in conjunction with the clarifications provided on their website FAQ.¹⁰ In accordance with the staging clarifications, scorers were instructed to use the F4-M1 derivation in M_{rec} and the C4-M1 derivation in M_{acc} as their primary channel for the identification of slow wave activity. PSGs were scored with Compumedics Profusion 3.4 (Build 353) software while viewed on Planar PX212M (1600 × 1200 resolution) LCD monitors via ATI Radeon HD2400 Pro 256 MB graphics cards. PSG scorers (CR, JM) had previously completed a laboratory training program for PSG scoring and participate regularly in intra- and inter-laboratory scoring concordance activities.

Statistical Analysis

Statistical analyses were performed using GraphPad Prism 6.02 (GraphPad Software, La Jolla, CA). The normality of paired data was determined by applying the Shapiro-Wilk test¹⁵ to the differences of the individual values to the group mean. Data are presented as mean and standard error or median and interquartile range where appropriate. Differences in sleep and arousal statistics between M_{rec} and M_{acc} were compared using paired t-tests. The mean difference and standard error of the difference between sleep and arousal statistics were also calculated to aid interpretation. Agreement between M_{rec} and M_{acc} for TST, EEG arousal index (Arl), and N3 proportions (% of TST) was determined according to the method of Bland and Altman.¹⁶ The first 10 PSGs were assessed for inter- and intra-scorer agreement. Epoch-by-epoch inter-scorer and intra-scorer agreement for sleep staging was determined by Cohen pairwise κ statistic.¹⁷ Inter-scorer and intra-scorer agreement for EEG arousals was assessed using proportion of specific agreement for positive ratings (PSA).¹⁸ The mean inter-scorer and intra-scorer agreement for each sleep stage and overall sleep was compared between M_{rec} vs. M_{acc} using paired t-tests. A box-and-whisker plot was also created to illustrate the dispersion

of inter-scorer and intra-scorer agreement for all sleep stages. A $p < 0.05$ was considered statistically significant.

RESULTS

Patients

Patient demographic data are presented in **Table 1**. Patients were middle-aged and obese. There was a male gender bias in this cohort, reflecting the reported gender-based prevalence in sleep disordered breathing. Obstructive sleep apnea was present in almost the entire cohort. Other sleep related diagnoses included: snoring and behaviorally insufficient sleep syndrome. One patient was deemed to have no sleep related disorder.

Sleep Statistics

PSG sleep statistics are presented in **Table 2**. There were no statistically significant differences between M_{rec} and M_{acc} for sleep and EEG arousal statistics. However, trends were evident toward an increased total sleep time, shorter sleep latency, and more sleep stage shifts when using M_{acc} .

Agreement Statistics

Scatter plots (**Figures 1, 3, 5**) and Bland-Altman plots (**Figures 2, 4, 6**) were created to determine the level of agreement between M_{rec} and M_{acc} for TST, ArI, and N3 proportions. Overall, these plots demonstrated relative parity between the 2 montages; however, there were substantial differences in a small number of patients. There was a bias for a higher mean (\pm SD)

TST of 4.9 min (\pm 28) when utilizing the M_{acc} compared to the M_{rec} montage. There was also a bias for higher mean (\pm SD) ArI of 1.1 (\pm 6.1) with the M_{acc} montage. No bias ($0.3\% \pm 6.2$) was detectable between M_{rec} and M_{acc} for N3 proportions.

Table 1—Patient characteristics of the study cohort

Number	50
Age, years	59 (45, 70)
BMI, kg/m ²	32.2 (29.2, 36.5)
Gender, M:F	36:14
Primary Sleep Diagnoses	
Obstructive sleep apnea	20
Mild	12
Moderate	12
Severe	1
Snoring	1
Behaviorally insufficient sleep disorder	1
No sleep disorder	1
AHI, /hour	16 (9.6, 30.5)
Comorbidities	
GERD	16
Hypertension	15
COPD	12
Type 2 diabetes	9
Ischemic heart disease	3

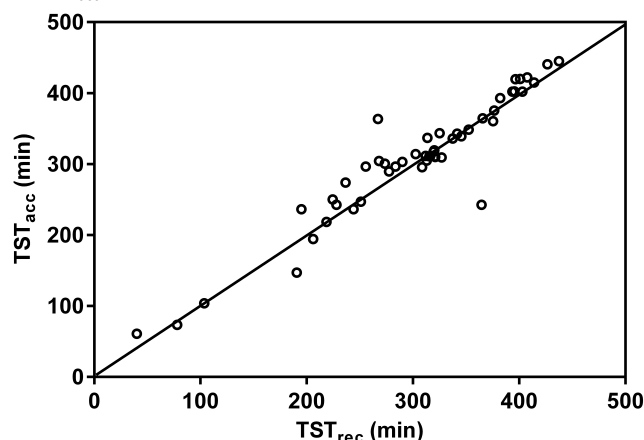
Values are presented as median (interquartile range). BMI; body mass index, AHI; apnea-hypopnea index, GERD; gastroesophageal reflux disease, COPD; chronic obstructive pulmonary disease. AHI was determined using the AASM alternative hypopnea definition.

Table 2—Calculated sleep indices using recommended and acceptable EEG montages

Sleep Parameter	Recommended	Acceptable	Difference	95% CI	p value
TST, min	303.2 (12.2)	308.1 (12.3)	-4.9 (6.5)	-22.8 – 13.0	0.589
Sleep latency, min	30.6 (6.0)	27.2 (5.0)	3.4 (6.5)	-9.7 – 16.6	0.606
REM latency, min	117.7 (11.1)	114.6 (10.9)	3.0 (8.9)	-14.8 – 20.9	0.730
Sleep efficiency, %	71.0 (2.8)	72.1 (2.8)	-1.1 (2.0)	-5.2 – 3.0	0.590
Sleep stage duration, min					
NREM Total	246.9 (9.2)	251.0 (9.9)	-3.4 (3.6)	-10.7 – 3.9	0.359
N1	34.4 (3.6)	37.7 (3.4)	-4.0 (3.7)	-3.4 – 11.5	0.278
N2	151.9 (7.7)	154.2 (8.2)	-2.3 (5.8)	-13.9 – 9.3	0.694
N3	60.7 (6.0)	59.8 (5.6)	0.9 (3.2)	-5.5 – 7.3	0.776
R	56.3 (4.6)	56.5 (4.8)	-0.2 (3.2)	-6.8 – 6.4	0.959
Sleep stage proportion, % of TST					
NREM Total	82.9 (1.2)	82.4 (1.3)	0.5 (1.1)	-1.8 – 2.8	0.679
N1	13.0 (1.5)	14.5 (1.8)	-1.5 (1.2)	-4.4 – 1.3	0.282
N2	50.6 (1.9)	48.9 (2.1)	1.7 (1.8)	-2.0 – 5.5	0.351
N3	19.3 (1.8)	19.0 (1.7)	0.3 (1.0)	-1.7 – 2.3	0.777
R	17.1 (1.2)	18.0 (1.4)	-0.9 (1.4)	-3.7 – 2.0	0.531
Stage shifts	160.9 (8.2)	172.3 (9.4)	-11.4 (6.7)	-24.7 – 2.0	0.094
EEG Arousals					
Total count	112.6 (9.6)	118.6 (10.2)	-6.0 (10.1)	-26.3 – 14.3	0.553
NREM count	94.5 (8.3)	99.4 (8.7)	-4.8 (8.8)	-22.5 – 12.8	0.586
REM count	18.0 (2.5)	19.3 (2.7)	-1.2 (2.0)	-5.2 – 2.7	0.539
Total Index, /h	24.3 (2.1)	25.4 (2.2)	-1.1 (1.8)	-4.6 – 2.4	0.531
NREM Index, /h	24.9 (2.2)	25.9 (2.4)	-1.0 (1.9)	-4.9 – 2.8	0.589
REM Index, /h	19.5 (2.1)	21.0 (2.2)	-1.5 (1.8)	-5.1 – 2.2	0.414

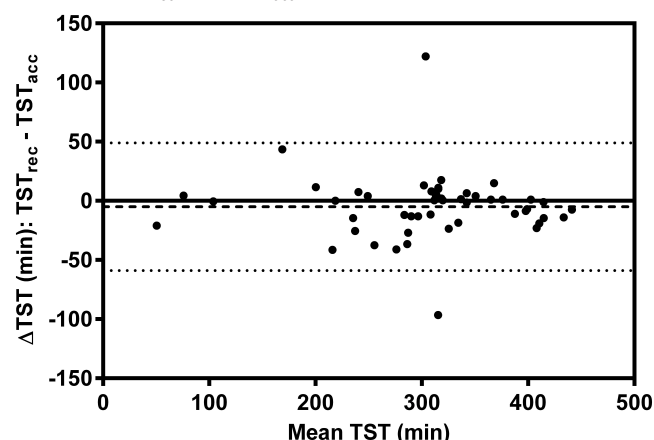
All values are presented as mean (SEM). TST, total sleep time. NREM; NREM sleep, N1; stage 1 sleep, N2; stage 2 sleep, N3; stage 3 sleep, R; REM sleep. p values represent paired t-test comparisons between M_{rec} and M_{acc} for each parameter.

Figure 1—Scatter plot comparing TST measured with M_{rec} and M_{acc} .



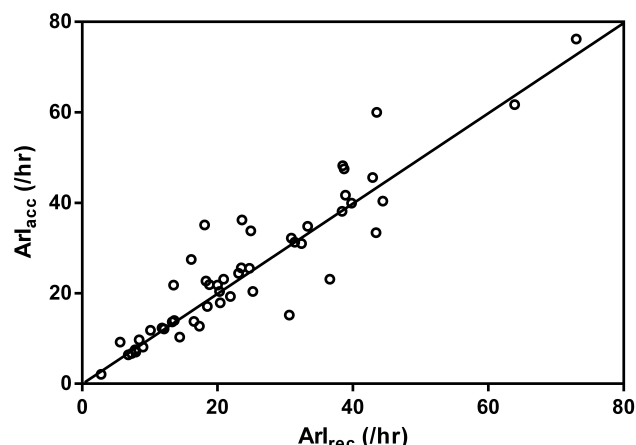
The diagonal line represents the line of identity. Correlation is significant at the 0.001 level.

Figure 2—Bland-Altman plot demonstrating the agreement between TST_{rec} and TST_{acc} .



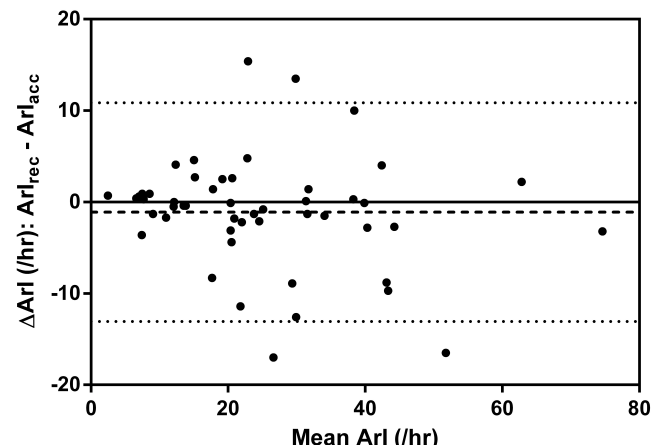
The dashed line represents the mean difference and the dotted lines represent the 5th and 95th percentile.

Figure 3—Scatter plot comparing the Arl measured with M_{rec} and M_{acc} .



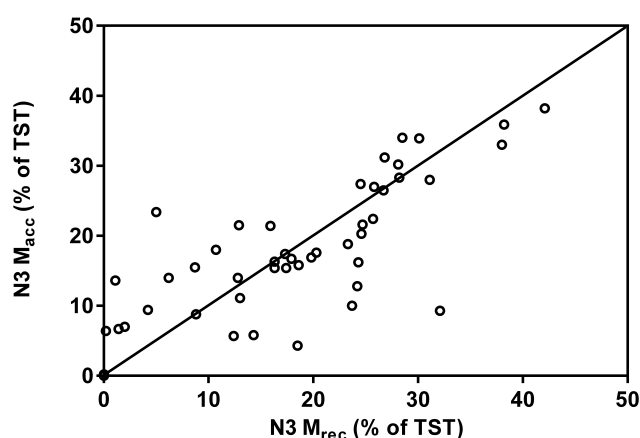
The diagonal line represents the line of identity. Correlation is significant at the 0.001 level.

Figure 4—Bland-Altman plot demonstrating the agreement between Arl_{rec} and Arl_{acc} .



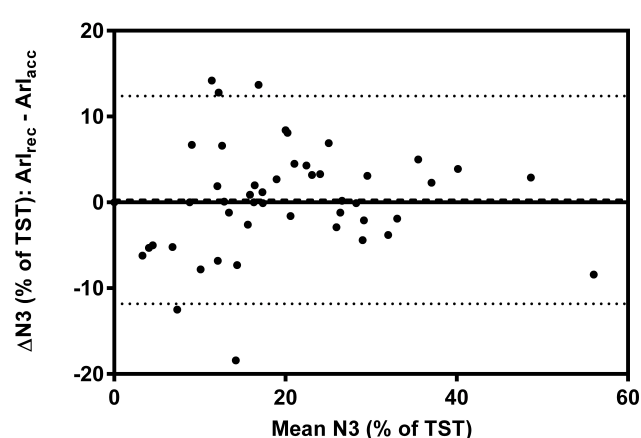
The dashed line represents the mean difference and the dotted lines represent the 5th and 95th percentile.

Figure 5—Scatter plot comparing stage N3 (as a proportion of TST) measured with M_{rec} and M_{acc} .



The diagonal line represents the line of identity. Correlation is significant at the 0.05 level.

Figure 6—Bland-Altman plot demonstrating the agreement of stage N3 (as a proportion of TST) between M_{rec} and M_{acc} .



The dashed line represents the mean difference and the dotted lines represent the 5th and 95th percentile.

Table 3—Inter-scorer agreement for polysomnograms scored using recommended and acceptable EEG montages

Parameter	Recommended	Acceptable	Difference	95% CI	p value
Sleep overall	0.74 (0.02)	0.61 (0.05)	0.12 (0.05)	0.02 – 0.23	0.047
Wake	0.88 (0.03)	0.81 (0.04)	0.07 (0.04)	-0.02 – 0.15	0.147
N1	0.47 (0.08)	0.42 (0.06)	0.05 (0.09)	-0.12 – 0.22	0.569
N2	0.68 (0.03)	0.55 (0.05)	0.13 (0.06)	0.02 – 0.24	0.053
N3	0.60 (0.08)	0.53 (0.09)	0.14 (0.12)	-0.10 – 0.37	0.281
R	0.92 (0.01)	0.83 (0.03)	0.09 (0.04)	0.01 – 0.16	0.042
EEG arousals	0.61 (0.09)	0.58 (0.07)	0.03 (0.07)	-0.12 – 0.17	0.369

Values expressed as mean (SEM). N1; stage 1 sleep, N2; stage 2 sleep, N3; stage 3 sleep, R; REM sleep. p values represent paired t-test comparisons between M_{rec} and M_{acc} for each parameter.

Table 4—Intra-scorer agreement for polysomnograms scored using recommended and acceptable EEG montages

Parameter	Recommended	Acceptable	Difference	95% CI	p value
Sleep overall	0.75 (0.03)	0.75 (0.02)	-0.01 (0.011)	-0.04 – 0.07	0.930
Wake	0.87 (0.03)	0.88 (0.02)	-0.01 (0.004)	-0.06 – 0.02	0.694
N1	0.51 (0.06)	0.48 (0.09)	0.03 (0.013)	-0.04 – 0.07	0.611
N2	0.66 (0.04)	0.68 (0.04)	-0.02 (0.009)	-0.12 – 0.04	0.835
N3	0.60 (0.08)	0.55 (0.10)	0.01 (0.016)	-0.10 – 0.08	0.874
R	0.92 (0.01)	0.87 (0.02)	0.05 (0.012)	0.01 – 0.14	0.421
EEG Arousals	0.61 (0.09)	0.57 (0.11)	0.03 (0.006)	-0.09 – 0.11	0.366

Values expressed as mean (SEM). N1; stage 1 sleep, N2; stage 2 sleep, N3; stage 3 sleep, R; REM sleep. p values represent paired t-test comparisons between M_{rec} and M_{acc} for each parameter.

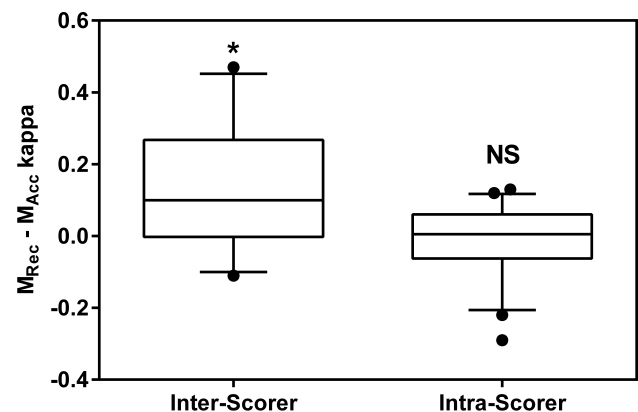
Inter- and Intra-Scorer Agreement

Inter-scorer and intra-scorer agreement data are presented in **Table 3**, **Table 4**, and **Figure 7**. Comparisons of epoch-by-epoch sleep staging between scorers demonstrated greater agreement in overall staging and REM sleep staging using M_{rec} . There was also a trend for greater agreement in staging N2 using M_{rec} . Comparisons of epoch-by-epoch sleep staging within scorers demonstrated no difference in overall staging agreement. There were no differences in the ability to discern EEG arousals using either montage.

DISCUSSION

To our knowledge, this is the first study to examine the differences between M_{rec} and M_{acc} with respect to sleep and EEG arousal statistics. In this study we found that the use of either M_{rec} or M_{acc} did not change the amount or proportions of sleep stages. Similarly, the ArI and the number of EEG arousals observed did not change appreciably between these montages. Despite the apparent equivalence between the two EEG derivations however M_{rec} was associated with greater inter-scorer agreement.

The primary aim of this study was to compare sleep and EEG arousal indices derived from the two derivations described in *The AASM Manual for the Scoring of Sleep and Associated Events* (M_{rec} and M_{acc}).⁴ Of these indices, we view TST as the principal sleep index in clinical PSG, as it not only indicates the adequacy of sleep during the PSG but is also used as a denominator for other key PSG statistics, such as the AHI and ArI. Our data suggest that M_{rec} and M_{acc} are similar with respect to TST. However, in two patients, we did find a large difference (> 1 h)

Figure 7—Box plots demonstrating paired inter- and intra-scorer agreement differences between M_{rec} and M_{acc} .

Agreement was calculated using Cohen's κ . Central line = median; Box boundary = 25th and 75th percentiles; Error bars = 10th and 90th percentiles; • = outliers. Differences were compared using paired t-tests. NS, not significant. * $p < 0.05$.

in the TST estimate between the montages. These large differences we attribute to a very high ArI in one patient and the degradation of the Cz signal in the other patient.

In this study we implemented the AASM FAQ V5 rule clarification for the staging of N3 when using M_{acc} . This made the C4-M1 derivation the primary channel for detection of slow wave activity in M_{acc} , while the primary channel in M_{rec} was the F4-M1 derivation. Two studies have already examined the differences between these two channels and demonstrated a statistically significant increase in N3 amounts and

proportions.^{19,20} The protocols of these studies differed from our study in that the study by Moser et al. examined AASM criteria compared to R&K,¹⁹ while the study by Ruehland et al. examined three EEG derivations versus one EEG derivation using AASM criteria for both derivations.²⁰ Despite the differences in study protocols, these two studies and our study are comparable with respect to comparisons of slow wave activity. Our finding of no difference in the amounts and proportions of N3 between M_{rec} and M_{acc} are therefore in contrast to these previously published studies.

There are a number of explanations for this difference in N3 amounts and proportions. The first explanation relates to the between-subject variations in slow wave amplitudes. While it has been well demonstrated that the F4-M1 derivation shows an average increase of 20% in slow wave amplitude compared to the C4-M1,¹³ this does not necessarily translate to a reduction in N3 sleep in all patients. The criteria for N3 requires only six seconds of slow wave activity with an amplitude of 75 μ V or greater across a 30-second epoch. This means that any slow waves with amplitudes greater than 90 μ V in the F4-M1 channel would still likely meet the amplitude criteria if observed in the C4-M1 channel.

Another explanation for the differences in N3 amounts and proportions could be related to laboratory approaches in discerning between K complexes and slow waves during sleep staging. Differentiating between K-complexes and slow waves can be difficult, especially when they are interspersed in the same epoch. This is not a new problem, as it was previously noted in the R&K as well as the Dement-Kleitman criteria.^{2,21} In this context, we are aware that some sleep laboratories will encourage their scorers to closely differentiate between K complexes and slow waves in the same epoch. Other sleep laboratories are more liberal in allowing the inclusion of K complexes (with > 75 μ V amplitude) that are interspersed with slow waves towards meeting the 20% epoch criteria for N3. This has the laboratory benefit of speeding up the staging process. Our laboratory prefers to closely differentiate between K-complexes and slow waves despite its inherent difficulties. While not being aware of how these previous studies have addressed such specific scoring subtleties, this phenomenon could be a factor in explaining the differing study results. Since the conduct of this study we note that the updated AASM Manual has addressed this issue by allowing the inclusion of K complexes in the determination of N3 criteria.²² This rule clarification may lead to further sleep stage agreement between laboratories.

The study by van Sweden et al. is the only other to have examined the differences in sleep staging statistics between a referential and bipolar EEG montage.²³ These authors compared the referential C4-M1 derivation with a bipolar montage of Fpz-Cz and Pz-Oz and described a tendency to stage greater amounts of slow wave sleep with the bipolar montage, while there were minimal differences in inter-scorer agreement. Although the increased slow wave sleep with the bipolar montage would seem to contradict the conventional wisdom of bipolar derivations and waveform amplitude, the inter-electrode distance of the Fpz-Cz derivation would closely approximate that of the F4-M1 referential derivation. This similarity of inter-electrode distance is likely the reason for the comparability of the bipolar

derivation with the C4-M1 derivation. Judging by the reference to van Sweden in the AASM supporting rationale paper,²⁴ this small study forms the basis for including the acceptable EEG montage into the AASM Manual,²⁴ though its common use at the Mayo Clinic may also have been a factor for inclusion.²⁶ While the AASM cited the van Sweden paper in its rationale, it recommended a frontal bipolar derivation with a smaller inter-electrode distance. Perhaps if the van Sweden bipolar derivations were used as part of M_{acc} , the FAQ V5 clarification would not be necessary.

The examination of inter-scorer agreement using Cohen κ demonstrated greater agreement when staging with M_{rec} . Part of this greater agreement may be attributed to the scorers' greater familiarity with M_{rec} . Since our laboratory routinely uses M_{rec} for clinical PSGs, it is likely that the scorers were perhaps more adept at identifying the different frequencies and morphological features in the recommended montage. Another factor that could explain greater agreement with M_{rec} is the confirmatory nature associated with the referential channels. The similarity of waveforms observed in both the F4-M1 and C4-M1 channels may lead to greater confidence in assigning a particular stage to an epoch by the scorers. We believe that with more training and familiarity with M_{acc} , it is likely that inter-scorer agreement would increase. When examining inter-scorer agreement for specific sleep stages, we observed statistically greater agreement in scoring the stage REM as well as a trend for better scoring of stage N2, although this was not statistically significant. This again may be associated with the scorers' familiarity with M_{rec} . The absolute agreement and hierarchy of staging agreement associated with each stage was consistent with the literature.^{20,26} Agreement was lowest in staging N1 sleep, better with stages N2 and N3, and highest with staging wakefulness and REM sleep.

The finding that the scoring of EEG arousals was similar for both EEG montages was not really surprising. Since the scoring of EEG arousals has quite poor agreement, it was unlikely that the number of PSGs used in this study would have enough power to find differences between the two EEG montages. Another contributing factor to this finding is the fact that rules for the scoring EEG arousals have not changed since the publication of the original EEG arousal standard in 1992.^{27,28} The comparison of EEG arousal scoring in this study is however limited by use of the PSA method. The PSA method may exaggerate agreement as it determines agreement for positive ratings but does not take into account negative ratings. Another problem associated with PSA method is that there is no specific duration of EEG arousal overlap criteria associated with this method. So even a 0.5-second overlap would signify EEG arousal agreement.

There are a few limitations associated with this study. The first limitation was that both scorers are from the same sleep laboratory. This aspect of our study has the potential to limit the generalizability of our study and could be addressed by the use of a number of scorers from different laboratories. The second limitation was the ability to properly blind the scorers. Even though all signals from the studies were labelled identically, it is likely that a number of factors affected the blinding. The first factor was the difference in appearance of the acceptable frontal channel (Fz-Cz) compared to the central channel

(C4-M1) during N3 was often readily apparent and commented on by the scorers. The second factor was the instruction to score all slow wave from the central (C4-M1) for M_{acc} likely alerted the scorers to these differences. In contrast we can point to the use of PSGs from a clinical cohort of subjects as strength of the study. Since all patients were studied for the suspicion of OSA, this increases the applicability of this study to other clinical sleep laboratories and their routine practice.

In conclusion, we found that sleep staging or EEG arousal statistics did not differ between the AASM recommended EEG montage or the acceptable EEG montage. An increase in inter-scorer agreement was associated with the use of M_{rec} , but this may be associated with the scorers' greater familiarity with this EEG montage. There appears to be no tangible advantage of using one EEG montage over the other in the clinical sleep laboratory setting.

ABBREVIATIONS

AASM, American Academy of Sleep Medicine
 AHI, apnea-hypopnea index
 ArI, EEG arousal index
 ECG, electrocardiogram
 EEG, electroencephalogram
 EOG, electrooculogram
 EMG, electromyogram
 FAQ, frequently asked questions
 LCD, liquid crystal display
 M_{rec} , montage using the AASM recommended EEG montage
 M_{acc} , montage using the AASM acceptable EEG montage
 NREM, NREM sleep
 N1, stage 1 sleep
 N2, stage 2 sleep
 N3, stage 3 sleep
 OSA, obstructive sleep apnea
 PSA, proportion of specific agreement
 PSG, polysomnography;
 R, REM sleep
 R&K, Rechtschaffen & Kales sleep staging criteria
 TST, total sleep time
 TST_{rec} , total sleep time determined with the AASM recommended EEG montage
 TST_{acc} , total sleep time determined with the AASM acceptable EEG montage

REFERENCES

1. Kushida CA, Littner MR, Morgenthaler T et al. Practice parameters for the indications for polysomnography and related procedures: an update for 2005. *Sleep* 2005;28:499-521.
2. Rechtschaffen A, Kales A. A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects. Los Angeles, CA: University of California, Brain Information Service/Brain Research Institute, 1968.
3. Himanen SL, Hasan J. Limitations of Rechtschaffen and Kales. *Sleep Med Rev* 2000;4:149-67.
4. Iber C, Ancoli-Israel S, Chesson A, Quan S, for the American Academy of Sleep Medicine. The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications. 1st ed. Westchester, IL: American Academy of Sleep Medicine, 2007.
5. Grigg-Damberger MM. The AASM scoring manual: a critical appraisal. *Curr Opin Pulmonary Med* 2009;15:540-9

6. McCormick L, Nielsen T, Nicolas A, et al. Topographical distribution of spindles and K-complexes in normal subjects. *Sleep* 1997;20:939-41.
7. Happe S, Anderer P, Gruber G, et al. Scalp topography of the spontaneous K-complex and of delta-waves in human sleep. *Brain Topogr* 2002;15:43-9.
8. Werth E, Achermann P, Borbely AA. Fronto-occipital EEG power gradients in human sleep. *J Sleep Res* 1997;6:102-12.
9. Adrian ED, Matthews BH. The Berger rhythm potential changes from the occipital lobes in man. *Brain* 1934;57:355-85.
10. American Academy of Sleep Medicine [Internet]. Scoring Manual FAQ. [cited 2013 December 4]. Available: <http://www.aasmnet.org/scoringmanualfaq.aspx>.
11. Epstein CM. The nuts and bolts of electroencephalography. *Sleep Med Clin* 2012;7:1-12.
12. Epstein CM, Brickley GP. Interelectrode distance and amplitude of the scalp EEG. *Electroencephalogr Clin Neurophysiol* 1985;60:287-92.
13. Kemp B, van Someren P, Roessen M et al. Assessment of human sleep depth is being de-standardized by recently advised EEG electrode locations. *PLoS One* 2013;8:e71234.
14. Urbaniak GC, Plous S. Research Randomizer (Version 3.0) [Computer software]. Retrieved on January 22, 2012, from <http://www.randomizer.org/>.
15. Shapiro SS, Wilk MB. An analysis of variance test for normality (complete sample). *Biometrika* 1965;52:591-611.
16. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307-10.
17. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960;20:37-46.
18. Fleiss JL. Measuring nominal scale agreement among many raters. *Psychol Bull* 1971;76:378-82.
19. 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.
20. 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.
21. Dement W, Kleitman N. Cyclical variations in EEG during sleep and their relation to eye movements, body motility and dreaming. *Clin Neurophysiol* 1957;9:673-90.
22. Berry RB, Brooks R, Gamaldo CE, et al. for the American Academy of Sleep Medicine. The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications, Version 2.0.3. www.aasmnet.org. Darien, IL: American Academy of Sleep Medicine, 2014.
23. van Sweden B, Kemp B, Kamphuisen HA, et al. Alternative electrode placement in (automatic) sleep scoring (Fpz-Cz/Pz-Oz versus C4-A1). *Sleep* 1990;13:279-83.
24. Silber MH, Ancoli-Israel S, Bonnet MH, et al. The visual scoring of sleep in adults. *J Clin Sleep Med* 2007;3:121-31.
25. Grigg-Damberger MM. The AASM Scoring Manual four years later. *J Clin Sleep Med* 2012;8:323-32.
26. 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.
27. EEG arousals: scoring rules and examples: a preliminary report from the Sleep Disorders Atlas Task Force of the American Sleep Disorders Association. *Sleep* 1992;15:173-84.
28. Bonnet MH, Doghramji K, Roehrs T, et al. The scoring of arousal in sleep: reliability, validity, and alternatives. *J Clin Sleep Med* 2007;3:133-45.

SUBMISSION & CORRESPONDENCE INFORMATION

Submitted for publication January, 2014

Submitted in final revised form February, 2014

Accepted for publication February, 2014

Address correspondence to: Brett Duce, Sleep Disorders Centre, Princess Alexandra Hospital, 199 Ipswich Rd, Woolloongabba, Brisbane, 4102, Queensland, Australia; Tel: +61 7 3176 5751; Fax: + 61 7 3176 7096; E-mail: brett.duce@health.qld.gov.au

DISCLOSURE STATEMENT

This was not an industry supported study. Mr. Duce has received honoraria from Philips-Respironics and Resmed. The other authors have indicated no financial conflicts of interest.