

ODDS RATIO PRODUCT OF SLEEP EEG AS A CONTINUOUS MEASURE OF SLEEP STATE

Odds Ratio Product of Sleep EEG as a Continuous Measure of Sleep State

Magdy Younes, MD^{1,2,3}; Michele Ostrowski, PRSGT¹; Marc Soiferman, MSc¹; Henry Younes, BSc¹; Mark Younes, BMu¹; Jill Raneri, PRSGT²; Patrick Hanly, MD²

¹YRT Ltd, Winnipeg, MB, Canada; ²Sleep Centre, Foothills Medical Centre, Calgary, Alberta, Canada; ³Sleep Disorders Centre, Winnipeg, Manitoba, Canada

Study Objectives: To develop and validate an algorithm that provides a continuous estimate of sleep depth from the electroencephalogram (EEG).

Design: Retrospective analysis of polysomnograms.

Setting: Research laboratory.

Participants: 114 patients who underwent clinical polysomnography in sleep centers at the University of Manitoba (n = 58) and the University of Calgary (n = 56).

Interventions: None.

Measurements and Results: Power spectrum of EEG was determined in 3-second epochs and divided into delta, theta, alpha-sigma, and beta frequency bands. The range of powers in each band was divided into 10 aliquots. EEG patterns were assigned a 4-digit number that reflects the relative power in the 4 frequency ranges (10,000 possible patterns). Probability of each pattern occurring in 30-s epochs staged awake was determined, resulting in a continuous probability value from 0% to 100%. This was divided by 40 (% of epochs staged awake) producing the odds ratio product (ORP), with a range of 0–2.5. In validation testing, average ORP decreased progressively as EEG progressed from wakefulness (2.19 ± 0.29) to stage N3 (0.13 ± 0.05). ORP < 1.0 predicted sleep and ORP > 2.0 predicted wakefulness in > 95% of 30-s epochs. Epochs with intermediate ORP occurred in unstable sleep with a high arousal index (> 70/h) and were subject to much interrater scoring variability. There was an excellent correlation ($r^2 = 0.98$) between ORP in current 30-s epochs and the likelihood of arousal or awakening occurring in the next 30-s epoch.

Conclusions: Our results support the use of the odds ratio product (ORP) as a continuous measure of sleep depth.

Keywords: sleep quality, sleep depth, sleep staging, ORP, odds ratio product

Citation: Younes M, Ostrowski M, Soiferman M, Younes H, Younes M, Raneri J, Hanly P. Odds ratio product of sleep EEG as a continuous measure of sleep state. *SLEEP* 2015;38(4):641–654.

INTRODUCTION

Since the inception of clinical polysomnography (PSG), evaluation of sleep and wakefulness states has followed the Rechtschaffen and Kales rules (R&K rules).¹ Sleep is divided into five discrete states, NREM stages 1, 2, 3, and 4, representing progressively deeper sleep, and REM sleep. Stages 3 and 4 NREM were recently combined as one stage (N3),² but, otherwise, there has been no change since 1968. R&K rules were formulated when sleep signals were recorded on paper and detailed analysis of the electroencephalogram (EEG) was not practical. It is widely recognized, however, that the division of sleep/wake states into distinct stages is artificial.^{3–6} A 30-s epoch is scored awake whether the EEG pattern is fully awake throughout or contains one or more brief sleep patterns (microsleep), so long as the duration of the sleep patterns is < 15 seconds. Once the 15-s threshold is exceeded, the stage suddenly becomes sleep. Likewise, sleep in any stage can be interrupted by EEG patterns associated with wakefulness but the whole epoch is scored as sleep, so long as the awake patterns are < 15 s in duration. Furthermore, the background EEG activity in the same R&K stage can differ substantially in visual appearance within and between epochs and patients. The EEG in stage 2 NREM (N2), for example, may be substantially similar to

stage 1 NREM (N1) (except for the occasional spindle or K complex) or similar to stage 3 NREM (N3) with delta waves just shy of 6 s per epoch.

The inability of R&K rules to capture the differences in EEG pattern between epochs of the same stage has been recognized for some time.^{3–6} With the advent of digital recordings and availability of sophisticated signal analysis techniques, temporal changes in the spectral EEG pattern as sleep evolves during a typical sleep cycle were extensively documented in the basic sleep literature. These studies showed that progression of NREM sleep towards delta sleep is typically associated with a gradual decrease in beta power and a gradual increase in delta power while sigma power initially rises then falls.^{3–6} These studies were directed primarily to understanding the basic mechanisms of sleep and a number of models have resulted from such studies.^{4,7–9}

The clinical significance of these gradual changes in the EEG is not known. Since arousal threshold increases progressively as sleep progresses from N1 to N3,^{10–14} it is reasonable to speculate that such a progression in spectral EEG features may reflect progression to deeper sleep. It would be useful to have a continuous, index of sleep depth that can be applied to the EEG in research and clinical studies performed to investigate complications of sleep disorders. Such an index may indicate, for example, that patients with the same amount of NREM stage 2 sleep may have very different average sleep depth, and this may account for different symptoms.

To our knowledge, there have been only two attempts at generating a continuous scale to evaluate sleep/wakefulness states during clinical polysomnography. Pardey et al.¹⁵ described an EEG analysis approach that utilizes a 10th order auto-regressive

Submitted for publication May, 2014

Submitted in final revised form September, 2014

Accepted for publication September, 2014

Address correspondence to: Magdy Younes, 1001 Wellington Crescent, Winnipeg, Manitoba, Canada R3M 0A7; Tel: (204) 297-0978; Fax: (866) 693-5084; Email: mkyounes@shaw.ca

model, followed by a multilayer perceptron (MLP) neural network, to generate a continuous scale between 100 (fully awake) and -100 (very deep NREM sleep). Performance of this model was tested in a study that compared the ability of the scale, vs. more conventional indices, to predict improvement in daytime sleepiness in obstructive sleep apnea patients following CPAP therapy.¹⁶ The new scale marginally improved the correlation when added to a multiple regression model, but only if the number of oxyhemoglobin saturation dips > 4% was excluded from the model. Asyali et al.¹⁷ proposed the use of the sum of alpha and beta powers as a continuous scale. There has been no validation of this scale beyond the original five files used in its development.

Concurrent with the basic studies on sleep mechanisms, digital analysis of the EEG extended into systems used for clinical polysomnography with the primary aim of replacing the inefficient manual scoring. Numerous digital approaches have been developed,^{18–25} and most commercial data acquisition systems include algorithms for automatic sleep scoring based on EEG time and frequency signal analysis. Most of these, however, aim to reproduce the manual scoring based on the R&K rules. Accordingly, they also fail to capture the differences in EEG patterns within and between epochs of the same stage.

In this paper we describe the development and validation of a continuous index for evaluating depth of sleep. It is an empiric scale that makes no a priori assumptions about what constitutes awake or sleeping EEG patterns. We show that: (a) the scale accurately predicts when the patient is awake, asleep or in an ambiguous state, (b) average scale decreases as NREM sleep progresses from N1 to N3, (c) the scale is highly variable within and between epochs and patients in the same R&K stage, and (d) the probability of an arousal or awakening increases progressively as the scale increases.

METHODS

Development of the Model

We used 58 anonymous PSG files previously recorded for unrelated research, approved by the Research Ethics Board of the University of Manitoba.^{14,26,27} The files, recorded in 2000–2001, included 2 central (C3/A2, C4/A1) and 1 occipital (O2/A1) EEG signals, 2 electro-oculograms, a chin electromyogram (EMG), an electrocardiogram, and signals from chest and abdomen Respiration (Respiration, Ambulatory Monitoring, Ardsley, NY, USA), thermistor flow, end-tidal PCO₂, oxyhemoglobin saturation (SpO₂), and a microphone. Signals were recorded at a sampling rate of 120 Hz using a Windaq data acquisition system (DATAQ Instruments, Akron, Ohio). The files were re-scored for sleep, arousals, respiratory events, and leg movements by an experienced PSG technologist (MO) who scored every 30-s epoch according to the 2007 American Academy of Sleep Medicine guidelines.²

A computer program, written in C#, analyzed the 2 central EEG signals in 3-s consecutive epochs. The power spectrum in each epoch was calculated in 0.33 Hz increments over the range 0.33 to 60.0 Hz, using fast Fourier transform. The sum of powers in the following 4 frequency ranges was calculated: 0.3–2.3 Hz (delta); 2.7–6.3 Hz (theta, excluding 6.7 and 7.0 Hz), 7.3–14.0 Hz (alpha-sigma), and 14.3–35.0 Hz (beta).

Frequencies 6.7 and 7.0 Hz were excluded from the theta range since slow alpha waves are often observed in some patients at these frequencies in bona fide NREM and REM sleep (alpha intrusion and REM alpha).

If a spindle is detected in a 3-s epoch and the alpha-sigma and/or beta powers were elevated, either or both powers were adjusted down to neutral levels depending on the dominant frequency in the spindle.

The files contained 415,924 3-s epochs. The delta powers in all these epochs were sorted and divided into 10 ranges with equal numbers (41,592 values per range). Ranks from 0 (lowest range) to 9 were assigned to the values in the different ranges. Theta, alpha-sigma, and beta powers were similarly processed resulting in 10 ranks for values in each frequency range. Each 3-s epoch was then assigned a 4-digit number (Bin #) representing, respectively, the ranks of delta, theta, alpha-sigma, and beta powers. Thus, a 3-s epoch with a Bin # 2815 indicates a power spectrum with low delta and alpha powers, high theta power, and moderate beta power. Accordingly, 10,000 different Bin numbers (0000 to 9999), reflecting 10,000 different power spectra, were theoretically possible.

The sleep state in which each 3-s epoch occurred was determined from the manual scoring. Three-second epochs falling during a 30-s epoch scored awake, or during an arousal in a 30-s epoch scored asleep (any stage), were designated “awake.” All other 3-s epochs were designated “asleep.” For each bin # we determined the probability of being “awake” from [total number designated “awake” / total number of epochs with the same bin # in the entire set] * 100. This resulted in a continuous probability range from 0 (never observed in awake epochs or during arousals) to 100 (never occurs during sleep periods). The number of 3-s epochs with a given bin # varied greatly and ranged from 6,802 occurrences (bin 9999) to 0 (1002 bin numbers were not represented). Bin numbers represented by < 10 samples in the entire set (3,781 bin numbers, accounting for 2.7% of all 3-s epochs) were assigned a neutral probability. The average number of epochs / bin # in the remaining bin numbers was 65 ± 137.

A look-up table was constructed that listed the probability of being “awake” for each of the 10,000 bin numbers. Although the probability could be expressed as a fraction or as a percent, we chose to divide the actual probability by 40, producing a range from 0 (always asleep) to 2.5 (always awake). The reason for dividing by 40 was that “awake” epochs in the development files represented 40% of all epochs, and it was felt at the time that a probability of 40% might be a suitable reference for the other values. This ratio was called the odds ratio product (ORP). Determination of this ratio is an integral first step for sleep staging in the recently introduced and validated²¹ automatic sleep scoring system.

Validation of the Odds-Ratio-Product as a Continuous Measure of Sleep-Wakefulness State

The objectives of this validation study were: (a) To confirm that the ORP values determined according to the approach outlined above, using the same look-up table, are predictive of sleep-wake states in other PSG files, recorded in a different institution, and scored by different technologists than the development files; (b) To determine whether 30-s epochs with

intermediate average ORP values, reflecting a mix of awake and sleep patterns, or basically unstable sleep, present scoring difficulty for technologists; (c) To determine the range of ORP values within and between epochs having the same R&K stage; (d) To determine if ORP values correlate with “arousability”; and (e) To determine the ORP values for different R&K stages in individuals without a sleep disorder. The validation study was approved by the Health Research Ethics Board at the University of Calgary.

The following patient groups were randomly selected from sleep center database at the University of Calgary to represent a broad spectrum of sleep pathology: severe OSA (AHI > 30, n = 8), moderate OSA (AHI 15–30, n = 10), mild OSA (AHI 5–15, n = 10), central sleep apnea (AHI > 15, n = 4), severe OSA on CPAP throughout (n = 5), periodic limb movement disorder (PLM index > 25, n = 4), insomnia (n = 5), narcolepsy (n = 5), no sleep pathology (n = 5). The PSG files contained the same variables as listed above for the development files but they were recorded with a Sandman system (Sandman, Tyco Healthcare, Kanata, Ontario, Canada) at a frequency of 128 Hz for the EEG electrodes. The PSG files were scored by 2 certified PSG technologists, each with over 10 years experience, one from the Sleep Centre at the University of Calgary and one from the Sleep Centre at the University of Manitoba. The technologist who scored the Development files was not one of the technologists who scored the Validation files. The files were also scored by the automatic system to generate the ORP values for the C3/A2 and C4/A1 EEG electrodes in consecutive 3-s epochs. The ORP values of the 2 electrodes were averaged, after excluding sections in which the signal from either electrode was corrupt (< 1% of all epochs).

Manually scored 30-s epochs were divided into 6 NREM stages and a single REM stage. The NREM stages were the conventional N1, N2, and N3, when both technologists agreed on the score and 3 stages where there was disagreement: W/N1, when one technologist scored the epoch as awake and the other as N1, N1/N2 for N1/N2 split, and N2/N3 for a split decision between N2 and N3. Epochs with N1/N3, REM/NREM, and W/REM disagreements were not analyzed, as they collectively represented < 2% of total epochs. Average ORP within each sleep stage, within all sleep epochs (total sleep time [TST]), and for the entire file were calculated.

Determination of the Relationship between ORP and Arousability in the Validation Files

Two approaches were used:

1) Determination of the probability of an arousal or awakening (A/AW) to occur in the next 30-s epoch when current ORP is within specified ranges (arousability index): For each file, 30-s epochs that were followed in the next epoch by a manually scored A/AW were identified. When the A/AW was scored by both technologists it was assigned a value of 1.0. When scored by only 1 of the 2 technologists, it was assigned a value of 0.5. The entire ORP range (0 to 2.5) was divided into 10 ranges, in 0.25 increments. For each ORP range we counted the number of epochs followed by an A/AW in the next epoch. A probability of A/AW occurring in the next epoch was calculated for each ORP range from [sum of subsequent A/AW in the ORP range * 100 / number of epochs in the ORP range].

This approach of using the ORP in the epoch preceding A/AW, as opposed to the epoch with A/AW, was to avoid biasing the ORP value by the A/AW being examined. This index simply indicates that if current ORP is “X” the likelihood of developing an A/AW soon after (within 30 sec) is “Y.” Total number of 30-s epochs in each ORP range was calculated for the entire set (56 files). Likewise, the total number of epochs, within each ORP range, that were followed by A/AW was calculated and an overall “arousability index” for each ORP range in the entire set was then calculated. The Pearson correlation coefficient was calculated to determine the relationship between average ORP and the arousability index.

2) Determination of the relationship between average ORP and the Arousal + Awakening Index (A/AW index) in different sleep stages: For each file we determined the average ORP within each of the NREM stages, excluding epochs scored as awake by one of the technologists (W/NR). We also determined the A/AW index from [total number of arousals and awakenings / time in hours] in each of the 5 stages. Stages represented by < 80 epochs in a given file were excluded since the confidence interval around the measured A/AW index with such numbers is quite large.²⁸ Because of the possible association between delta power and arousability,^{13,14} we also calculated average delta power and average log delta power in each stage. Multiple regression analysis was performed with the A/AW index as the dependent variable and average ORP and average delta power as the independent variables.

RESULTS

Observations from the Development Files

The patients whose files were used for development (42M/16F) were 20 to 80 years old (51.1 ± 12.9), with a body mass index of $31.6 \pm 5.8 \text{ kg/m}^2$. Total sleep time was 262 ± 77 min, and sleep efficiency (SE) was $63\% \pm 17\%$. Arousal index ranged from 9 to 95/h (33 ± 21 /h). Thirteen patients had mild obstructive sleep apnea (OSA) (AHI 5–15/h), 17 had moderate OSA (AHI 15–30/h), 17 had severe OSA (AHI > 30/h, average AHI 61 ± 33), 5 had insomnia (SE $56\% \pm 10\%$), and 6 had no apparent sleep pathology. Thirty of the 34 patients with moderate-severe OSA received CPAP in the second half of the study, and many of these had normal sleep during this time.

The probability of a given Bin # falling in an awake 30-s epoch or during arousal ranged from 0% to 100%. The correlation coefficient between % awake and the corresponding ranks in the different frequency ranges was highest for beta rank ($r = 0.65$) and theta rank ($r = -0.61$) while correlations with delta ($r = -0.29$) and alpha ($r = 0.22$) ranks were considerably lower. In multiple linear regression (MLR) theta rank had the highest coefficient (-7.7% per rank; $P = 0$), followed by beta rank (6.5% per rank; $P = 0$) and alpha rank (2.6% per rank; $P = 0$). Delta rank had no significant impact by MLR (coefficient = -0.02% per rank, $P = 0.86$).

Figure 1A shows the interaction between the 2 main determinants (theta and beta ranks) on % awake. It is clear that the likelihood of wakefulness with any theta rank is highly dependent on the beta rank, and vice versa, and neither rank can predict wakefulness on its own.

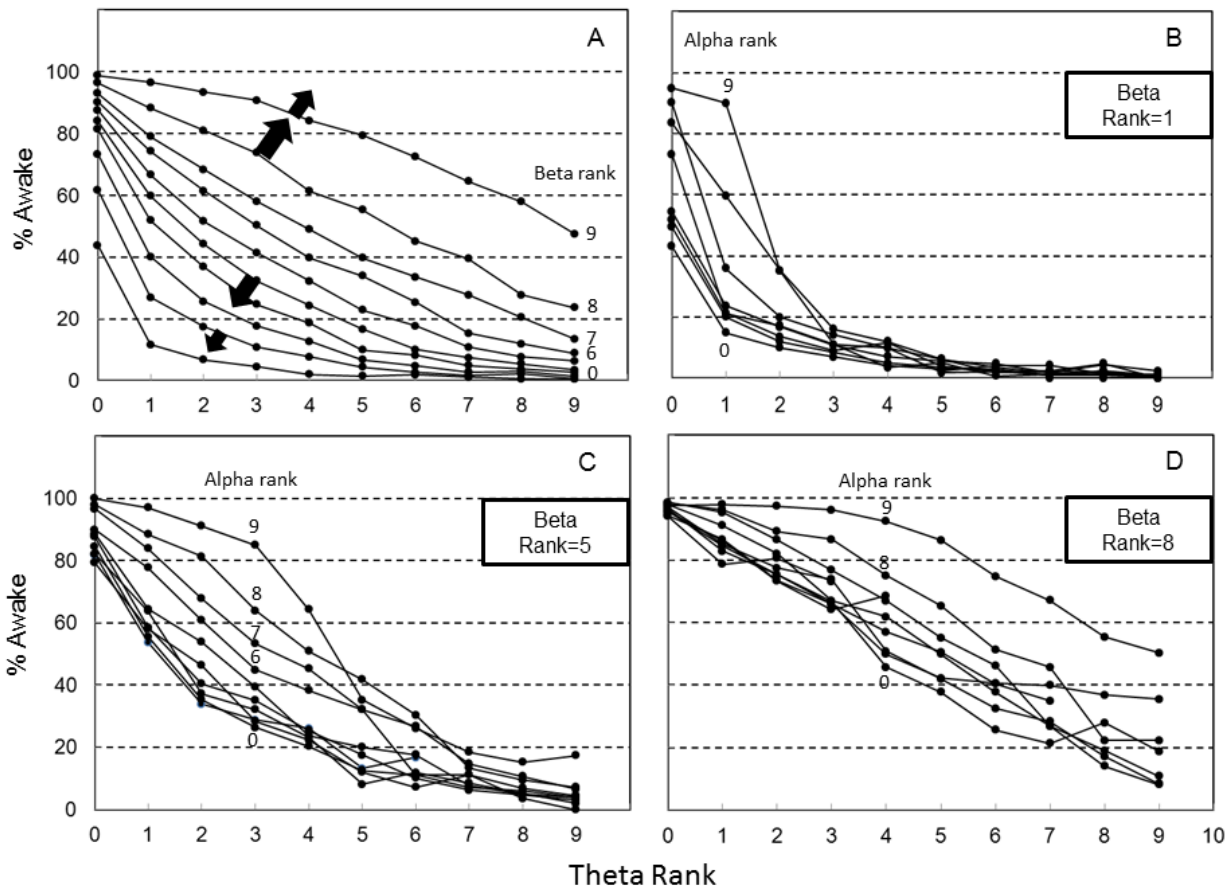


Figure 1—Percent of 3-s epochs falling within 30-s epochs staged manually as awake (% Awake) as a function of relative power in different frequency ranges. (A) % Awake decreases as theta power increases but the relationship is markedly affected by the concurrent beta power. A high delta power increases % Awake when % Awake is high according to theta/beta powers and decreases it when theta/beta powers are in a low % Awake (Block arrows). (B,C) Impact of alpha power on % awake at different theta-beta combinations. Note that the same % Awake is observed at highly varied EEG spectral patterns.

The effect of alpha rank was quite complex and dependent on the concurrent theta and beta ranks. Figures 1B to 1D show the probability of wakefulness as a function of theta power at three different beta ranks (rank 1, rank 5, and rank 8). The effect of alpha rank was minimal in all cases at alpha ranks < 5 (note the crowding of lines at alpha levels ≤ 5 in Figures 1B to 1D). At low beta levels (Figure 1B), alpha rank had a marked effect on % awake only when theta rank was low. At beta rank of 5 (Figure 1C), alpha rank increased the probability of wakefulness in the theta range 1 to 6 while at high beta levels (Rank 8, Figure 1D) the alpha effect was prominent at all but the lowest theta levels.

Whereas the effects of beta and alpha ranks on the probability were always positive, and the effect of theta rank was always negative, the effect of delta rank was more complex and bi-directional depending on the other ranks. In general, when the awake probability was $< 40\%$ based on the other 3 ranks, a higher delta rank caused a further reduction in likelihood of the epoch being scored awake. On the other hand, when the other three ranks resulted in an awake probability $> 60\%$, a higher delta rank caused a further increase in the likelihood of the epoch being scored awake. Thus, a higher delta rank expanded the perimeter of the theta/beta relation (Block arrows, Figure 1A).

The average of the 10 ORP values in each 30-s epoch in the development files ($n = 41,952$) was calculated. An average

ORP value of 1.5 distinguished between awake and asleep 30-s epochs with an accuracy of 94.8%.

Ability of ORP Values to Predict Sleep/Wake States in Validation Files

Patient characteristics and clinical diagnoses for patients in the validation study are shown in Table 1. There was a total of 44,274 30-s epochs. Figure 2 shows the frequency of occurrence of 30-s epochs with different ranges of average ORP (i.e., average of the ten 3-s epochs within each 30-s epoch). Within each bar, solid and white sections represent the fractions scored as awake and asleep, respectively, by both scorers. Hatched sections represent epochs scored awake by one scorer and asleep by the other. Frequency was highest for epochs with the lowest average ORP and lowest for epochs with intermediate ORP. The fraction of epochs staged asleep by both scorers (white sections) decreased progressively as average ORP increased, from 99% when average ORP was < 0.25 to 0% when ORP was > 2.25 . Concurrently, the fraction of epochs scored as awake by both scorers increased progressively from 0.1% at the lowest ORP to 98% at the highest. Importantly, the fraction that received a split score increased from a minimum value of 1.1% with the lowest ORP to 27.5% when average ORP was between 1.25 and 1.50 and then decreased thereafter to 1.7% at the highest ORP.

Table 1—Patient characteristics.

Diagnosis	#	Age	Gender	BMI (kg/m ²)	TST (min)	SE (%)	AHI (/h)	Arousal Index (/h)	PLM Index (/h)
No Pathology	5	49.6 (16.6)	1M/4F	29.0 (6.1)	360 (69)	85 (9)	2 (2)	13 (4)	0 (0)
Narcolepsy	5	35.0 (13.3)	2M/3F	25.4 (4.5)	366 (36)	82 (6)	5 (1)	16 (5)	2 (2)
Insomnia	5	51.6 (6.6)	1M/4F	43.6 (19.9)	123 (79)	29 (18)	4 (2)	21 (15)	5 (10)
PLM Disorder	4	56.3 (4.1)	2M/2F	34.5 (12.9)	286 (88)	69 (24)	4 (2)	27 (6)	43 (36)
Mild OSA	12*	45.1 (17.2)	2M/10F	31.1 (5.4)	294 (74)	71 (17)	8 (3)	27 (8)	15 (22)
Moderate OSA	9**	52.6 (12.8)	6M/3F	35.9 (6.8)	235 (89)	67 (17)	24 (4)	26 (11)	20 (47)
Severe OSA	7***	53.4 (6.7)	3M/4F	43.3 (16.5)	126 (43)	60 (26)	61 (31)	43 (24)	27 (40)
Central Apnea	4*	59.4 (7.9)	3M/1F	31.4 (5.8)	225 (114)	77 (11)	61 (21)	26 (9)	21 (32)
OHS	5**	60.2 (9.0)	2M/3F	43.5 (9.7)	182 (54)	79 (10)	28 (16)	17 (8)	22 (37)

For patients who received a split study the values given are for the pre-CPAP period. Values in parenthesis are standard deviation. *Three of these patients were on CPAP throughout and three had a PLM index > 45. **One of these patients was continuously on CPAP, another had a PLM index of 144, and four underwent a split study. ***All patients in this group underwent a split study, and two had a PLM index > 45. *Three of these patients underwent a split study, and one had a PLM index of 68. **One of these patients was continuously on CPAP, one had a PLM index of 87, and four underwent a split study. AHI, apnea hypopnea index; BMI, body mass index; OHS, obesity hypoventilation syndrome; PLM, periodic limb movements; SE, sleep efficiency; TST, total sleep time.

The data of Figure 2 indicate that when the average ORP of a 30-s epoch is ≤ 1.0 the epoch was scored asleep by at least one scorer in 98% of cases and by both scorers in 94% of cases. When average ORP is ≥ 2.0, the epoch was scored awake by at least one scorer in 99% of cases and by both scorers in 95% of cases. If an ORP of 1.5 is used as a cutoff between awake and asleep, the overall accuracy was 94% for agreement with at least one scorer and 86% for agreement with both scorers.

Ranges of ORP Between and Within Stages and Patients

Figure 3 is a representative example of the changes in ORP observed over a 3-h period in one patient. The patient progressed quickly after lights-out into stage N1 and then N2 (top panel) and remained in N2 for most of the 3 hours except for occasional awakenings, and brief changes in R&K stage to N1 or N3. During the awake periods, ORP was between 2.0 and 2.5. ORP decreased gradually upon falling asleep from each awake period. When the duration of continuous sleep was sufficiently long, ORP decreased to below 0.5, at which time N3 was observed. The rate of decline in ORP following awakenings varied considerably (Figure 3, top). The top panel also shows when AASM arousals were scored (+ symbols). Arousals essentially disappeared once ORP approached 0.5.

The 6 tracings in Figure 3 represent six 30-s epochs from the same patient. The top 2 were both scored awake, but the lower one (Tracing B) included a 12-s period of sleep (bar) during which ORP transiently decreased. Tracing C was scored by both scorers as N1. Tracings D and E were both scored as N2, with tracing D looking much like N1 except for the presence of spindles, while tracing E looked much like N3 (Tracing F) except that the duration of delta waves fell short of 6 seconds. The average ORP values in these 6 epochs are shown by the arrows in the ORP panel. Average ORP decreased progressively from 2.4 in tracing A to 0.33 and 0.32 in tracings E and F.

Table 2 shows the average (± SD) number of epochs in different R&K stages. During scoring, 7.9% ± 6.6% of epochs received a split awake/NREM decision, and more epochs

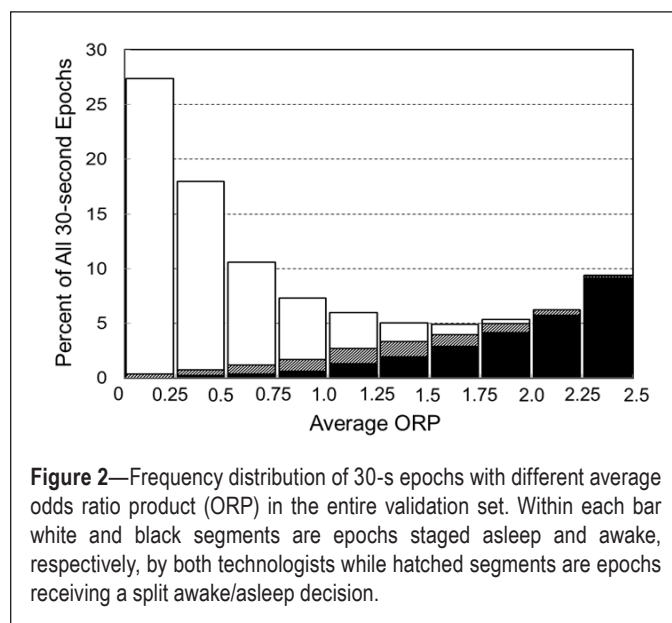


Figure 2—Frequency distribution of 30-s epochs with different average odds ratio product (ORP) in the entire validation set. Within each bar white and black segments are epochs staged asleep and awake, respectively, by both technologists while hatched segments are epochs receiving a split awake/asleep decision.

received a split N2/N3 decision (8.8 ± 8.2) than a consensus N3 decision (5.3 ± 6.5).

Figure 4 shows average ORP in different R&K sleep stages for individual patients and the overall average (± SD) in each stage. Each dot is the average of all ORP values calculated during all epochs in the same stage/patient. Average ORP (open circles) decreased progressively from R&K stage awake (W2, Figure 4, ORP = 1.88 ± 0.35) to the deepest sleep (N3; ORP = 0.13 ± 0.05). Epochs in which the scorers were split between 2 stages (e.g., W/N1, N1/N2...etc.) had an average ORP that was intermediate between the 2 adjacent stages. Average ORP in REM sleep was comparable to that in stage N1. The variability in average ORP among patients was evident in all stages, especially in REM sleep. It should be noted, however, that the variability in N3 was very small among patients who reached this stage (n = 36), with average ORP falling below 0.30 in all patients.

As seen in Figure 4, average ORP during R&K stage awake (W2) was < 2.0 in many patients (30/56). This low average may be due to the presence of microsleep in many epochs manually scored awake (e.g., Figure 2B), to the presence of artifacts (e.g.,

blink artifacts in the EEG) that would increase theta power, thereby artificially reducing ORP, or to an atypical EEG pattern during wakefulness. To distinguish between these possibilities, we obtained the average of the highest 5 ORP values in each 30-s epoch scored awake, on the grounds that these must represent clearly an awake pattern (the epoch would not have been scored awake otherwise). The data under stage W1 in Figure 4 represent the results. The average here was considerably higher than for W2 (2.19 ± 0.29 vs. 1.88 ± 0.35), indicating that the lower average ORP in W2 was to a considerable extent due to transient events (most commonly microsleep) within the awake epochs. Nonetheless, in 12 patients average ORP remained below 2.0 in W1, including one patient in whom ORP was < 1.0 (Figure 4). Inspection of the awake epochs in these patients revealed 4 reasons that accounted for the low ORP. These are shown in Figure 5 and described below:

In 5 patients, waves that appeared to be alpha waves with the normal screen compression (1.0 cm/sec) were in reality in the theta range (6.0 ± 0.7 Hz). Figure 5A is an example from a patient with a dominant awake frequency of 5.7 Hz (the

Table 2—Distribution of sleep stages.

Stage	Number of Patients	# Epochs/Patient Mean (SD)	% of Total Epochs Mean (SD)
Awake	56	214 (166)	27.0 (19.6)
Awake/NREM	56	64 (53)	7.9 (6.6)
NREM-1	56	38 (29)	4.6 (3.1)
NREM-1/2	56	55 (37)	6.7 (3.9)
NREM-2	56	230 (128)	28.1 (13.1)
NREM-2/3	52	64 (54)	8.8 (8.2)
NREM-3	36	40 (50)	5.3 (6.5)
REM	51	67 (49)	8.8 (5.8)

SD, standard deviation.

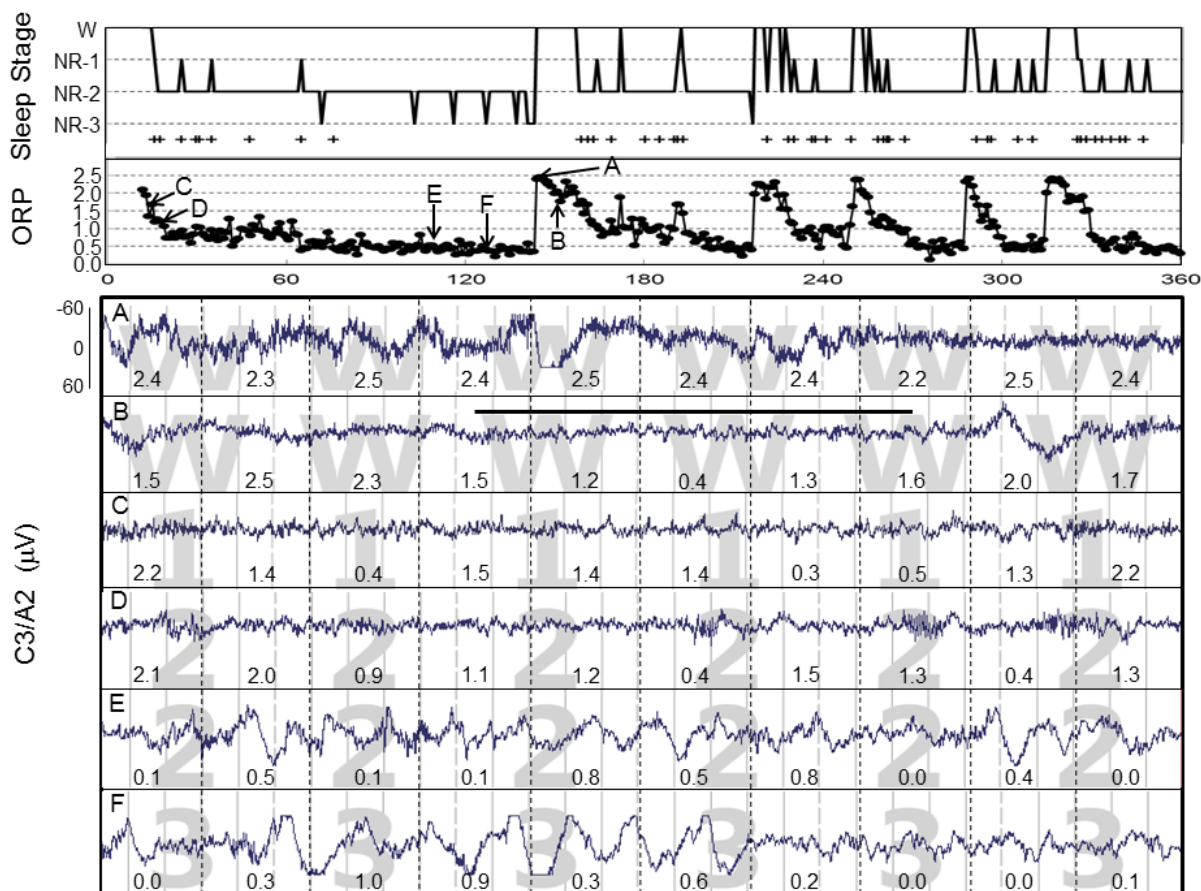


Figure 3—Top panel: Sleep histogram from a representative patient spanning the first 3 hours of a study. Abscissa, epoch #. Arousals are indicated with a plus sign (+). The second panel gives the corresponding odds ratio product (ORP) values. Each dot is a 30-s epoch. Note that ORP rises to near 2.5 with every awakening and then declines gradually. Note also that stage N3 began appearing as ORP approached 0.5 and arousals disappeared simultaneously. A to F are the points from which the tracings below were obtained. Tracings A and B were staged awake, but in A the patient was fully awake, while in B he was having short periods of sleep. Note the marked difference in ORP distribution within the two awake panels and the transient reduction in ORP during the unstable state of tracing B. Tracing C was staged N1 but it is visually substantially similar to B and has a similar average ORP. D and E were both staged N2. Their visual appearances are quite different, with D looking more like N1, but with spindles, while E looks more like N3 (F). ORP values are quite different between the two N2 epochs. Panel F was staged N3. ORP at this point was 0.34 while it was similar (0.33) in the previous tracing (E), which was staged N2. The difference was only that delta waves exceeded 6 seconds in F but not in E.

one with the lowest ORP). These waves greatly increased the measured theta power during awake epochs forcing the ORP down (Figure 1A). When present, the pattern of a dominant theta frequency during wakefulness was consistent throughout the file. In 3 patients, the EEG during awake epochs was indistinguishable from that in stages N1 or REM (Figure 5B). As recommended by the AASM guidelines (AASM) in such cases (absence of dominant alpha/beta rhythm), stage awake was scored because of other criteria (eye movements with high chin EMG). In 3 patients the EEG had extremely low power in all frequency ranges (Bins 0000, 1000, 1010). These bin numbers are assigned a low ORP in the look-up table because they frequently occur also in stages N1 and REM, albeit not continuously throughout the 30-s epoch. Finally, in one patient the average of the 2 central electrodes was artificially reduced because of a peculiar artifact in one of the electrodes (Figure 5D) that was not picked up by any of the multiple algorithms for detecting bad EEG signals.

It is worth noting that these 12 patients with atypical awake EEG contributed 65% of epochs that were staged awake by both scorers while having an average ORP < 1.0. Without them, epochs with ORP < 1.0 would predict sleep in 99.3% of cases. As may be expected, these 12 patients had very few epochs with an average ORP > 2.0; namely $2.9\% \pm 2.3\%$ of epochs. This is significantly lower than in patients with typical awake EEG, where the fraction of epochs with ORP > 2.0 was 19.6 ± 17.8 ($P < 0.002$). The high ORP epochs in patients with

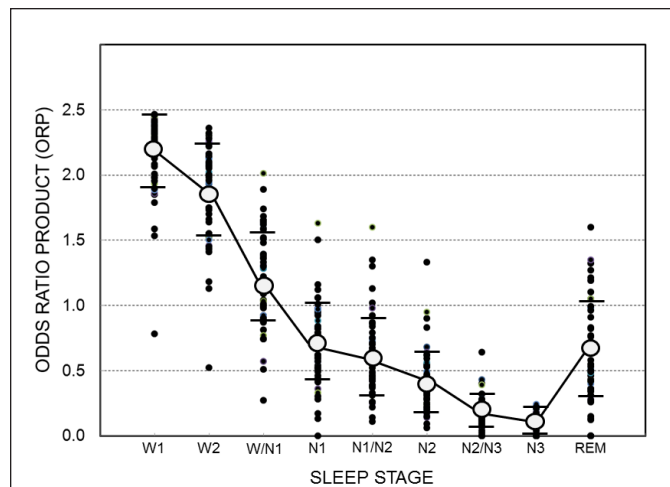


Figure 4—Average ORP values in different sleep stages. Each solid dot is a different patient. Open circles and horizontal bars are means \pm standard deviation. W2, N1, N2, N3, and REM are standard Rechtschaffen and Kales stages, with W2 representing awake epochs. W1 is average of the highest 5 ORP values within 30-s epochs staged awake (see text). W/N1, N1/N2 and N2/N3 are epochs that received a split score by the 2 technologists. $n = 56$ for W1 to N2, 52 in N2/N3, 36 in N3 and 51 in REM. Note that average ORP in N3 was < 0.3 in all patients who reached that stage and that many patients had comparable ORP values in lower stages (N2/N3, N2, N1/N2 and even N1). Also note the large variability among patients in ORP during REM sleep.

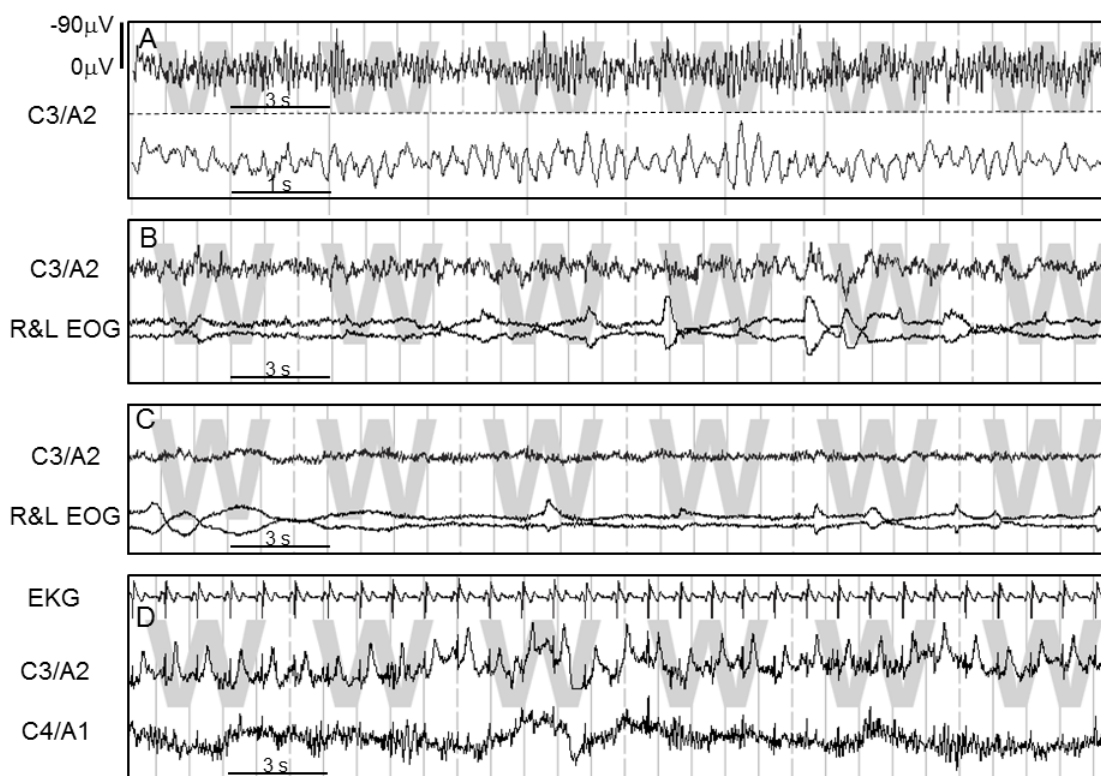


Figure 5—Four examples of atypical awake EEG patterns that cause the odds ratio product (ORP) to be underestimated during wakefulness. C3/A2 and C4/A1 are the two central electrodes. R&L EOG, are the eye electrodes. EKG, electrocardiogram. Calibration of EEG is shown in the top panel and is the same in all panels. (A) Predominant wave frequency (6.0 Hz) is in the theta and not the alpha range. (B) EEG is very similar to N1 but there are rapid eye movements and high chin EMG activity (not shown). (C) Very low power EEG. (D) A very large T wave cardiac artifact in C3/A2 not removed by the EKG removal algorithm because of its unusual pattern and distance from the R wave, resulting in high theta power in one electrode.

Table 3—Probability of an arousal or an awakening occurring in the next 30-second epoch in different current ORP ranges.

Current ORP Range	Total # of Epochs	# Files with Data in ORP Range	Mean (Range) of Epoch #'s per File	Total # of Arousals in Next Epoch	Total # of Awakening in Next Epoch
0.00–0.25	10,388	55	189 (0–580)	655	188
0.25–0.50	8,321	56	149 (14–363)	1,003	286
0.50–0.75	5,298	56	95 (19–225)	938	331
0.75–1.00	3,258	56	58 (4–181)	683	326
1.00–1.25	2,236	54	41 (0–146)	551	268
1.25–1.50	1,484	48	31 (0–151)	389	202
1.50–1.75	900	35	25 (0–108)	264	118
1.75–2.00	573	26	22 (0–88)	200	100
2.00–2.25	219	15	15 (0–123)	74	51

ORP, odds ratio product.

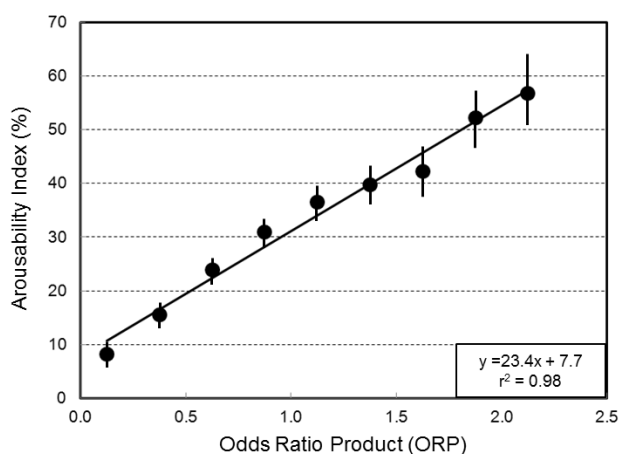


Figure 6—Relationship between average ORP in current 30-s epochs and the likelihood of an arousal or awakening occurring in the next 30-s epoch (arousability index). Vertical bars are the confidence interval of the probability.

atypical EEG occurred only when beta power was very high to compensate for the high theta power during wakefulness (Figure 1).

Relationship between ORP and Arousability

Table 3 shows the total number of 30-s epochs scored as asleep by at least one technologist, grouped in different ORP ranges. The table also gives the numbers of arousals and awakenings occurring in the following epoch in each ORP range. The total number of epochs available for analysis was several thousand up to the fifth ORP range (1.0–1.25), and nearly all files contributed epochs in this range. The total number of available epochs decreased rapidly at higher ORP levels as progressively more epochs were scored awake (Figure 2).

Figure 6 shows the overall arousability index (sum of subsequent arousals and awakenings * 100 / total number of epochs) as a function of current ORP. Epochs in the 0 to 0.25 ORP range were followed by A/AW in only 8.1% of cases. This ratio increased as ORP increased, reaching 56% at the highest ORP range. There was a near perfect correlation between current average ORP and average arousability index. The confidence

interval around the index (fraction) was extremely narrow at low ORP levels and increased gradually as number of available epochs decreased.

Figure 7A shows the association between average ORP in specified R&K stages and the A/AW index in the same epochs. Each point is the average ORP and the corresponding A/AW index calculated from all 30-s epochs in a given stage (symbols in Figure 7B) in one patient. With very few exceptions (see below) A/AW index was < 20/h when average ORP was < 0.4, regardless of whether R&K stage was N2, N3, or in between. A/AW index was very high whenever average ORP was > 1.0. A/AW index was quite variable in the intermediate ORP range. The average relation was sigmoid (dashed line, Figure 7A). Nonetheless, there was an excellent linear relationship between the 2 variables (solid line, Figure 7A; $r = 0.86$, $P < 1E-34$). Since delta power is also believed to reflect arousability,^{13,14} we plotted the association between delta power and A/AW index for the same data (Figure 7B). There appeared to be a breakpoint at delta power of $\approx 200 \mu V^2$. The entire range of A/AW index occurred below this level while at levels > 200 μV^2 A/AW index was quite low, regardless of R&K stage, and showed only a minimal dependence on delta power over a very wide range. Using Log delta power instead of average power did not improve the correlation. Nonetheless, there was a significant linear relationship ($r = 0.41$, $P < 0.0001$) between delta power and A/AW index. Following multiple linear regression (MLR) analysis with A/AW index as the dependent variable and ORP and delta power as the independent variables, delta power had no significant effect.

We investigated the reasons for the variable relationship between A/AW index and ORP in the ORP range 0.3 to 1.0 by examining the EEG records of data where the average A/AW index was low relative to average ORP (right arrows, Figure 7A), and of those with the opposite relationship (left arrows, Figure 7A). In the former case (too few arousals) the reason was invariably that a large proportion of epochs in the scored (in the figure) R&K stage consisted of a mix of brief awake and light sleep patterns, accounting for the high ORP, with the awake patterns not contributing 15 s to justify a score of “awake” while not meeting the AASM criteria for scoring arousal because they were too brief or not preceded by 10 s of stable sleep.² Figure 8, top, is an example from the data set

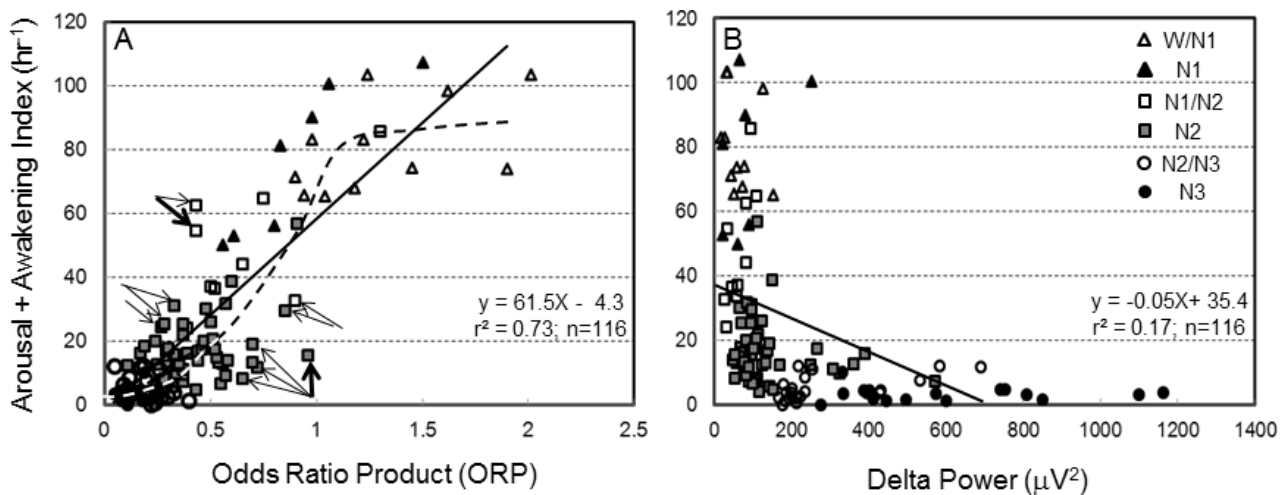


Figure 7—(A) Relationship between average ORP in different sleep stages (see legend in panel B) and the corresponding arousal/awakening index. Each dot is the average ORP in all epochs having the same sleep stage in one patient. Note the highly variable arousal/awakening index among patients in stages N1, N2 and N1/N2. Straight line, linear regression line (see inset for regression data). Dashed line is moving average showing a sigmoid relation between ORP and the index. Arrows indicate outliers that were examined in detail (Figure 8). **(B)** Relationship between delta power and arousal + awakening index.

indicated by the heavy right arrow in Figure 7A (ORP = 0.96, A/AW index = 15).

In cases where there were too many arousals relative to average ORP (left arrows, Figure 7A) the problem was invariably excessive intrusion of waves with a dominant frequency in the 6.0 to 7.0 Hz range. An example is shown in Figure 8, bottom. In calculating ORP, power in frequencies 6.7 and 7.0 Hz are deliberately not included in either the alpha or the theta power calculation (see Methods). Accordingly, when there is excessive power in this borderline frequency range some of the power is included in the theta range while very little is added to the alpha range. Thus ORP is deliberately forced down (see Discussion).

ORP Values in Patients with Different Disorders

Although the study was not designed to evaluate differences in ORP among various sleep disorders, it was of interest to see how these disorders differed in this limited sample. Figure 9 shows the average individual ORP values in different stages, in total sleep time and in total recording time in the various disorders. It is clear that, except for N3, individual ORP values varied widely within each stage in patients with all disorders represented here, and that there was much overlap in the ORP values between different disorders, and between patients with and without a recognizable sleep disorder. As seen in the total sleep time panel, average ORP for all sleep periods, which takes into account time spent in different stages and ORP within each stage, ranged between 0.1 and 1.6. For total recording time, which takes sleep efficiency and awake ORP into account as well, average ORP ranged from 0.2 (very deep sleep) to 2.2 (nearly awake continuously).

DISCUSSION

In this communication we have introduced a new, continuous index for evaluating the level sleep depth (ORP), with a range extending from full wakefulness to the deepest sleep. We have presented evidence that this index does in fact reflect

depth/quality of sleep. In the process of developing and validating this index we have found that: (1) The ORP range (0 to 2.5) can be divided into three ranges, with range 0 to 1.0 predicting sleep and range 2.0 to 2.5 predicting wakefulness with > 95% accuracy in both cases, while range 1.0 to 2.0 represents unstable sleep. (2) ORP varies considerably in the same R&K stage within and between individuals, and there is much overlap in ORP values between different R&K stages. (3) Depth/quality of sleep is determined primarily by the balance between powers in a pro-sleep theta frequency range and a pro-awake beta frequency, with less important contributions from powers in the alpha and delta frequencies. (4) ORP values among patients with the same sleep disorder cover a wide range, suggesting that the effect of the disorder on sleep quality is highly variable and is often negligible. We believe that determination of this index may be a useful addition to in-laboratory and home sleep studies.

Evidence that ORP Reflects Sleep Quality/Depth

This evidence derives from two sets of data. First, the pattern of full wakefulness, recognized by prominence of alpha and/or beta waves,^{1,2} can readily be considered as the state of highest vigilance in polysomnography data, while stage N3 is recognized as the deepest sleep with the highest arousal threshold.¹⁰⁻¹⁴ ORP was found to progressively decrease, on average, from full wakefulness (W1, Figure 4) to stage N3 (Figure 4). With few exceptions in which awake EEG was atypical (see below), average ORP during full wakefulness was > 2.0. When the awake EEG pattern was typical, with alpha/beta prominence ($n = 44$), there was little variability between patients (ORP = 2.31 ± 0.11). More importantly, whenever a patient reached N3, average ORP was very low and there was little variability between patients (0.13 ± 0.05). These observations indicate not only that ORP is correlated with sleep depth, but also that ORP is substantially unaffected by technical or individual factors. With respect to technical factors, it is worth emphasizing that ORP calculations used in the validation study

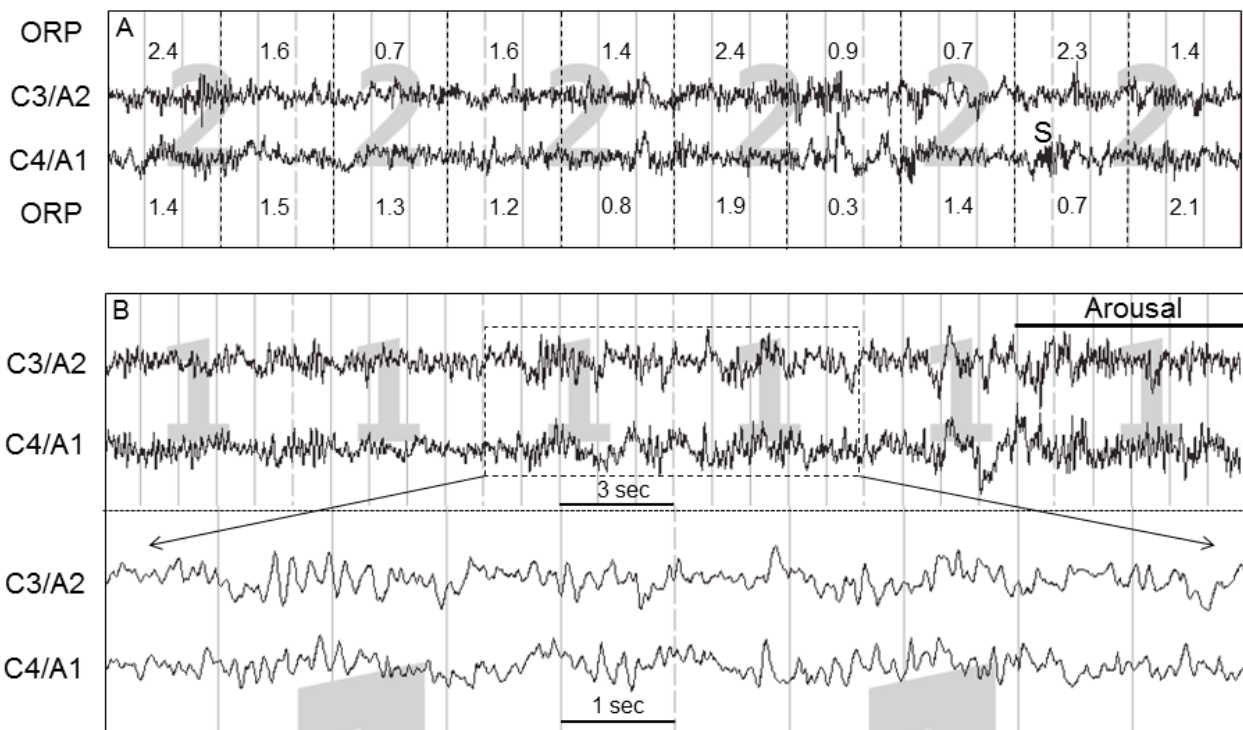


Figure 8—(A) tracing showing an example from a patient in whom arousal and awakening index was low relative to ORP in stage N2 (heavy right arrow in Figure 7). ORP, odds ratio product. C3/A2 and C4/A1 are 2 central EEG electrodes. Note the frequent alpha/beta intrusions that did not meet arousal criteria because they were too brief or were not preceded by 10 s of continuous sleep. ORP was high, reflecting these intrusions. (B) Tracing showing an example from a patient in whom arousal and awakening index was high relative to ORP (heavy left arrow in Figure 7). Note the extensive intrusion by waves that appear like alpha waves but are in reality in the theta range (predominant frequency = 6.3 Hz). This causes underestimation of ORP.

were based on data from files used in the development study, which were recorded in a different institution by a completely different data acquisition system, and a look-up table that was based on the scoring of these development files by a different technologist. Furthermore, to the extent that the patients were of both genders and covered a wide range of age, body habitus, and sleep disorders, the fact that in the deepest sleep and in full typical wakefulness ORP varies little between patients indicates that ORP is not sensitive to differences in EEG attributes between patients.

The insensitivity of this index to technical and individual factors stems from our use of an empiric approach to estimating the ORP value. Thus, the % awake (ORP) value assigned to any given EEG pattern (bin #) is based on data from several hundred thousand 3-s epochs that likely included all conceivable patterns that occur in awake epochs. When a certain pattern is found exclusively in awake epochs, it can be safely assumed that it represents an awake pattern. Likewise when a pattern is never found in awake periods, it can be safely assumed that it is a sleep pattern. More importantly, awake or sleep patterns were not restricted to specific EEG patterns based on a priori assumptions of what constitutes sleep or wakefulness. Each of the 10,000 possible patterns was tested for its likelihood of occurrence in epochs staged as awake or asleep. As a result, the same ORP can be reached with highly varied EEG patterns (% Awake, Figure 1). This approach identified hundreds of patterns (bin numbers) that occur almost exclusively during wakefulness, and vice versa.

The second line of evidence to validate ORP as a measure of sleep quality/depth is the association between ORP and arousability. It can be safely assumed that a sleep state from which it is more difficult to arouse is deeper than a state from which one can be easily aroused. When data from all 56 patients were pooled together, we found that the ORP in a given current 30-s epoch was almost perfectly correlated with the likelihood of an arousal or awakening occurring in the next epoch (Figure 6). Figure 7A also shows that when average ORP in epochs belonging to a given R&K stage is > 1.0 , A/AW index is invariably very high. Conversely, when average ORP is < 0.4 , A/AW index is almost invariably very low. In the “in-between” ORP range A/AW index was quite variable. But, a considerable part of this variability has to do with arousal scoring guidelines and the occurrence of much power in the low-alpha/high-theta frequency range in some files (Figure 8). We do not wish to imply that ORP is a direct measure of arousal threshold in each individual patient. Clearly, differences in the strength of arousal stimuli among patients as well as inter-individual differences in the relationship between ORP and arousal threshold will influence the arousability at a given ORP in individual patients.

Why Should a Probability-of-Being-Awake Index Reflect Depth of Sleep?

Although the correlations between ORP and conventionally accepted sleep depth (Figure 4) and arousability (Figures 6 and 7) are clear, the reason why an index of the probability of an EEG pattern occurring in awake epochs would reflect sleep

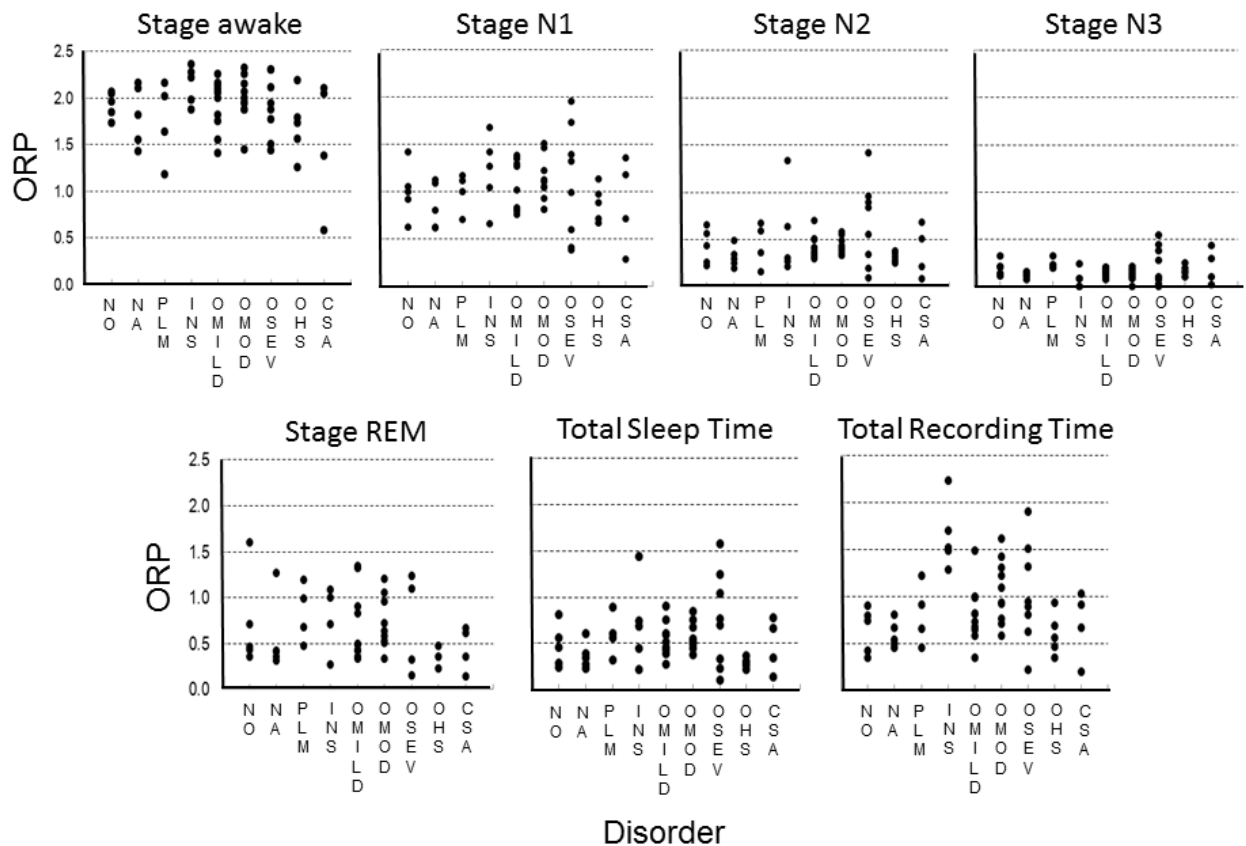


Figure 9—Odds ratio product (ORP) in different sleep stages and for total sleep time and total recording time in patients with no identified pathology during the sleep study (NO) and in patients with various disorders. NA, narcolepsy; PLM, periodic limb movement disorder; INS, insomnia; OMILD, OMOD, and OSEV are patients with mild, moderate and severe obstructive sleep apnea; OHS, obesity hypoventilation syndrome; CSA, central sleep apnea. Note the wide range of ORP within each group and the considerable overlap between groups. See Table 1.

depth during sleep may not be readily apparent. We arbitrarily divided the 3-s EEG frequency spectrum into a maximum of 10,000 patterns and found that at least 90% of these were encountered at least once in the EEG signals of the development data set. Six thousand patterns were encountered more than 10 times, and 877 patterns were encountered more than 100 times. Evidently, the 3-s EEG spectral pattern is extremely heterogeneous. Each of these patterns was found to have a specific probability of occurring during a 30-s epoch staged as awake. A 3-s pattern that is found in awake epochs 90% of the time (ORP = 2.25) not only indicates that the pattern is most likely an awake pattern but also that if the epoch included sleep patterns these were not long enough to qualify the epoch to be staged asleep (i.e., micro-sleep). These brief sleep periods within awake epochs clearly represent the lightest of sleep; they spontaneously revert to an awake pattern within seconds. When the same pattern occurs in a 30-s epoch staged as asleep, it indicates that the epoch contains periods of very light sleep or arousal. Conversely, a 3-s pattern that is found only 10% of the time during 30-s epochs staged awake (ORP = 0.25) not only indicates that the patient is most likely asleep but also that the 30-s epoch in which the pattern occurred is very unlikely to contain a 15-s period of awake patterns to justify manual staging of awake. This very small likelihood of awakening from this epoch indicates a deeper sleep. By extension,

a pattern with a 50% likelihood of occurring in awake periods (ORP = 1.25) might indicate that this pattern tends to occur in unstable sleep, with a 50% chance of the 30-s epoch being staged awake. The average ORP in a 30-s epoch is the net of ORP values in the ten 3-s values. A high average value indicates preponderance of light sleep patterns, and vice versa.

What Constitutes Deep Sleep?

One important implication of the excellent correlations between ORP and arousability/depth of sleep (Figures 6 and 7) is that there is no unique EEG pattern that reflects sleep depth. As shown in Figure 1, a deep sleep pattern (low % awake or ORP) can take many forms. For example an ORP that reflects deep sleep (e.g., 10% awake or ORP of 0.25) is found with several combinations of theta and beta powers that even include high beta power (Figures 1C and 1D). Thus, progression into deep sleep does not follow a regimented pattern, such as a progressive increase in delta power or a progressive decrease in alpha/beta power. In fact we found that while a high delta power during sleep is associated with very low arousability, a low delta power can also be associated with low arousability (Figure 7B). Also, alpha/beta power need not decrease as sleep deepens. The power in these high frequencies may stay unchanged or even increase so long as theta power increases more (Figure 1). Our findings, therefore, emphasize the importance

of theta power in determining sleep depth. This was not emphasized in previous models.^{4,7-9}

The highly varied patterns that are associated with deep sleep are extremely difficult to distinguish by eye from other patterns associated with light sleep, particularly because of the widely held belief that sleep depth is related to delta power. The visual differences may be very subtle but our data clearly show that these subtle differences reflect very different sleep depths (Figures 4, 6, and 7).

Potential Applications

1) Attended polysomnography: Figures 4, 7, and 9 clearly show that ORP varies over a wide range within the same R&K stage. To the extent that ORP reflects sleep depth and arousability, calculation of average ORP during individual stages, total sleep time and total recording time could provide an added dimension to the evaluation of patient's sleep quality. This additional information may help explain symptoms that might not be explained by the conventional sleep report. It could also be used in research that explores the relationship between sleep quality and various cognitive, metabolic, and psychiatric disorders, an area that is currently growing in importance.²⁹⁻³⁹

It should be pointed out that average ORP in total sleep time (TST) incorporates the effect of relative times spent in different sleep stages. Thus, a patient who spends much time in stage N1 will have a high ORP in TST even if ORP within each sleep stage is "normal" for that stage. Furthermore, average ORP in total recording time incorporates the effect of sleep efficiency. This index (ORP in total recording time) therefore incorporates all the factors currently considered to reflect sleep quality, including arousal/awakening index (which would increase ORP during a given sleep stage), distribution of times in different sleep stages, and sleep efficiency.

Including average ORP in 30-s epochs may also help improve the quality of manual or automatic sleep scoring. Figure 2 shows that 30-s epochs with an ORP value between 1.0 and 2.0 account for 65% of all epochs that resulted in a split awake/asleep decision between the two technologists. At the same time, our data show that ORP values in this range represent very poor sleep (Figures 6 and 7). Forcing technologists to make decisions about the status of these epochs not only results in confusing inter-rater variability but also, depending on the technologist, may result in scoring many epochs as stage N2 when sleep quality is in fact very poor. It may be reasonable, therefore, to not attempt to give a definite R&K stage to these epochs and simply stage them as unstable sleep.

2) Level 3 sleep studies: The use of level 3 studies for home diagnosis of respiratory sleep disorders is rapidly increasing. These studies lack a signal that would indicate if the patient was asleep during monitoring, and this can cause underestimation of OSA severity.^{40,41} Two of the main advantages of ORP are that it can be derived from one EEG electrode and it does not require expensive, time-consuming manual scoring. Requirement of multiple electrode placement and manual scoring are the main deterrents to sleep monitoring in the home. The current results were obtained from individual C electrodes. The correlation between the 30-s ORPs obtained from C3 and C4 averaged 0.89 ± 0.11 in the 56 validation studies, so that

either electrode can suffice. There is no reason why the ORP method cannot be adapted to frontal electrodes (see limitations, below) to make it easier for patients to apply the electrodes and the hardware required to process the signal and generate ORP can be extremely noninvasive.

Our current data indicate that an ORP value < 1.0 indicates sleep and a value > 2.0 indicates wakefulness with 95% accuracy. The less common ORPs in the 1.0 to 2.0 range can simply be considered unstable sleep. Such information could greatly enhance the utility of level 3 studies without much additional technical requirements or cost.

3) Free-standing monitor of sleep quality: A large number of patients have nonspecific complaints that could be related to poor sleep, such as nonrestorative sleep, excessive sleepiness, fibromyalgia, fatigue, and depression. In many cases, the history does not suggest one of the more recognizable sleep disorders, such as sleep apnea or PLMs, to justify a sleep study. Having a simple home test that evaluates duration and quality of sleep may help identify those who have an underlying problem with their sleep.

4) Real-time indicator of the state of vigilance and depth of sedation: Although EEG in the current study was performed after the sleep studies were completed, the algorithms that extract the ORP values can easily output the results in real-time, every three seconds. This, coupled with the need for only one electrode, would make it possible to monitor the state of vigilance during awake states that require a high level of vigilance, or to control sedation and anesthesia in intensive care units and operating rooms. Further studies are needed before the utility of ORP in these applications can be established.

Limitations

The main limitation is the occurrence in some patients of atypical awake EEG patterns, and the rare presence of theta-frequency artifacts, during wakefulness (Figure 5). These may cause ORP to be quite low, (W2, Figure 4) and if a threshold of 2.0 is used to identify wakefulness, awake periods may be missed. Obviously this error would occur only when eye and EMG electrodes are not available to indicate wakefulness. This problem can be overcome to some extent by counting the highest ORPs in the 30-s epochs when applying the threshold (compare W1 and W2, Figure 4). However, a few cases may still remain (W1, Figure 4). The significance of having such atypical EEG patterns, with no real alpha activity, during wakefulness is not known, and it is possible that they may reflect some pathology. This remains to be determined. These types of atypical awake EEG patterns can likely be identified through more detailed analysis of power at specific frequencies within the theta range, and by implementing additional algorithms to detect specific artifacts (e.g., Figure 5D), and to distinguish between the pattern of extremely low total power (bin# 0000, Figure 5C) when it occurs during wakefulness and during stages N1 and REM sleep. It is our experience that when this pattern is encountered during wakefulness it is continuous throughout the 30-s epoch (Figure 5C), whereas it lasts for only several seconds when it occurs during sleep.

We had hoped to identify normal ORP values from the data of patients with no identifiable sleep pathology. However, we found nearly complete overlap between the results in these

patients and those from patients with sleep disorders (Figure 9). It is possible that some of the apparently healthy subjects were not entirely normal since they were all referred for evaluation of a suspected sleep disorder. At present, our results suggest that an ORP > 1.0 represents very poor sleep (Figure 7A). Because ORP values of < 0.3 were achieved in stages N3 and N2/N3 in nearly all patients and average ORP of < 0.5 was reached for total sleep time (TST) in many patients and even for total recording time in some patients (Figure 9), it is reasonable to state that a value < 0.3 for any sleep stage and of 0.5 for TST represent very good sleep quality. More studies that are designed primarily to measure ORP values in normal subjects from different genders and age groups are needed to address this issue.

Arousability was inferred here from the likelihood of occurrence of arousals in epochs with different ORP values. It would be of interest and importance to determine whether this measure of arousability correlates with the arousal threshold measured from cortical response to specific stimulations.

Our results were based on analysis of central EEG data. Use of central EEG electrodes in some of the potential applications listed above may present some technical difficulty. Studies are needed to determine if analysis of frontal EEG data with the current algorithms and look-up table yield similar results. If not, a separate look-up table would have to be developed for frontal leads.

The PSG records used for the validation study did not include frontal leads. It is possible that some epochs in N3 (as judged by the frontal leads, if available) may have been manually scored as N2 if delta wave amplitude was borderline in the central electrodes. The ORP in such epochs would invariably be very low, but the low ORP would be credited to N2 instead of N3, thereby possibly accounting for some of the very low ORP values in N2 (e.g. Figure 4).

Finally, this algorithm cannot distinguish between REM and NREM sleep on its own. EOG electrodes would be needed if this distinction is required.

In summary, our results support the use of the odds ratio product (ORP) both to distinguish sleep and wakefulness and as a continuous measure of sleep depth. This measurement can be considered as an adjunct or even an alternative to conventional sleep analysis. Further studies are required to evaluate how well the change in ORP compares to other clinical outcome measures.

DISCLOSURE STATEMENT

This study was supported by YRT Ltd, Winnipeg, Manitoba, Canada. Dr. Younes is majority owner of YRT Ltd. Ms. Ostrowski, Mr. Soiferman, and Mark Younes are currently employed by YRT. Henry Younes is minority owner of YRT. The other authors have indicated no financial conflicts of interest.

REFERENCES

1. Rechtschaffen A, Kales A. A manual of standardized terminology, techniques, and scoring system for sleep stages of human subjects. Washington, DC: U.S. Government Printing Office, 1968. NIH Publication No. 204.
2. Iber C, Ancoli-Israel S, Chesson AL Jr., Quan SF. The AASM manual for the scoring of sleep and associated events. Westchester, IL: American Academy of Sleep Medicine, 2007.

3. Knott JR, Gibbs FA, Henry CE. Fourier transforms of the electroencephalogram during sleep. *J Exp Psychol* 1942;31:465–77.
4. Helli M, Fortune RD. The neuronal transition probability (NTP) model for the dynamic progression of non-REM sleep EEG: the role of the suprachiasmatic nucleus. *PLoS One* 2011;6:e23593.
5. Uchida S, Maloney T, Feinberg I. Beta (20–28 Hz) and Delta (0.3–3 Hz) EEGs oscillate reciprocally across NREM and REM Sleep. *Sleep* 1992;15:352–8.
6. Uchida S, Maloney T, March JD, Azari R, Feinberg I. Sigma (12–15 Hz) and Delta (0.3–3 Hz) EEG oscillate reciprocally within NREM sleep. *Brain Res Bull* 1991;27:93–6.
7. Borbely AA. A two process model of sleep regulation. *Hum Neurobiol* 1982;1:195–204.
8. Borbely AA, Achermann P. Sleep homeostasis and models of sleep regulation. *J Biol Rhythms* 1999;14:559–70.
9. Achermann P, Borbely AA. Sleep homeostasis and models of sleep regulation. In: Kryger MH, Roth T, Dement WC, eds. Principles and practice of sleep medicine. 5th ed. Philadelphia, PA: WB Saunders, 2011.
10. Williams HL, Hammack JT, Daly RL, Dement WC, Lubin A. Responses to auditory stimulation, sleep loss, and the EEG stages of sleep. *Electroencephalogr Clin Neurophysiol* 1964;16:269–79.
11. Berry RB, Bonnet MH, Light RW. Effect of ethanol on the arousal response to airway occlusion during sleep in normal subjects. *Am Rev Respir Dis* 1992;145:445–52.
12. Gugger M, Molloy J, Gould GA, et al. Ventilatory and arousal responses to added inspiratory resistance during sleep. *Am Rev Respir Dis* 1989;140:1301–7.
13. Berry RB, Asyali MA, McNellis MI, Khoo MC. Within-night variation in respiratory effort preceding apnea termination and EEG delta power in sleep apnea. *J Appl Physiol* 1998;85:1434–41.
14. Younes M. Role of arousals in the pathogenesis of obstructive sleep apnea. *Am J Respir Crit Care Med* 2004;169:623–33.
15. Pardey J, Roberts S, Tarassenko L, Stradling J. A new approach to the analysis of the human sleep/wakefulness continuum. *J Sleep Res* 1996;5:201–10.
16. Bennett LS, Langford BA, Stradling JR, Davies RJ. Sleep fragmentation indices as predictors of daytime sleepiness and nCPAP response in obstructive sleep apnea. *Am J Respir Crit Care Med* 1998;158:778–86.
17. Asyali MH, Berry RB, Khoo MC, Altinok A. Determining a continuous marker for sleep depth. *Comput Biol Med* 2007;37:1600–9.
18. Pittman SD, MacDonald MM, Fogel RB, et al. Assessment of automated scoring of polysomnographic recordings in a population with suspected sleep-disordered breathing. *Sleep* 2004;27:1394–403.
19. 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 24 × 7 utilizing the Siesta database. *Neuropsychobiology* 2005;51:115–33.
20. Svetnik V, Ma J, Soper KA, et al. Evaluation of automated and semiautomated scoring of polysomnographic recordings from a clinical trial using zolpidem in the treatment of insomnia. *Sleep* 2007;30:1562–74.
21. Malhotra A, Younes M, Kuna ST, et al. Performance of an automated polysomnography scoring system versus computer-assisted manual scoring. *Sleep* 2013;36:573–82.
22. Ronzhina M, Janoušek O, Kolářová J, Nováková M, Honzík P, Provazník I. Sleep scoring using artificial neural networks. *Sleep Med Rev* 2012;16:251–63.
23. Stepnowsky C, Levendowski D, Popovic D, Ayappa I, Rapoport DM. Scoring accuracy of automated sleep staging from a bipolar electroocular recording compared to manual scoring by multiple raters. *Sleep Med* 2013;14:1199–207.
24. Virkkala J, Hasan J, Värrä A, Himanen SL, Müller K. Automatic sleep stage classification using two-channel electro-oculograph. *J Neurosci Methods* 2007;166:109–15.
25. Shambroom JR, Fábregas SE, Johnstone J. Validation of an automated wireless system to monitor sleep in healthy adults. *J Sleep Res* 2012;21:221–30.
26. Younes M, Ostrowski M, Thompson W, Leslie C, Shewchuk W. Chemical control stability in patients with obstructive sleep apnea. *Am J Respir Crit Care Med* 2001;163:1181–90.
27. Younes M. Contributions of upper airway mechanics and control mechanisms to severity of obstructive apnea. *Am J Respir Crit Care Med* 2003;168:645–58.

28. Goldstein A, ed. *Biostatistics: an introductory text*. New York: The MacMillan Company, 1964.
29. Colten HR, Altevogt BM, eds. *Committee on sleep medicine and research, sleep disorders and sleep deprivation: an unmet public health problem*. Washington, DC: The National Academies Press, 2006.
30. Paudel M, Taylor BC, Ancoli-Israel S, et al. Sleep disturbances and risk of depression in older men. *Sleep* 2013;36:1033–40.
31. Blackwell T, Yaffe K, Ancoli-Israel S, et al. Osteoporotic Fractures in Men (MrOS) Study Group. Association of sleep characteristics and cognition in older community-dwelling men: the MrOS sleep study. *Sleep* 2011;34:1347–56.
32. Cohen-Zion M, Stepnowsky C, Marler, Shochat T, Kripke DF, Ancoli-Israel S. Changes in cognitive function associated with sleep disordered breathing in older people. *J Am Geriatr Soc* 2001;49:1622–7.
33. Lim AS, Kowgier M, Yu L, Buchman AS, Bennett DA. Sleep fragmentation and the risk of incident Alzheimer's disease and cognitive decline in older persons. *Sleep* 2013;36:1027–32.
34. Xie L, Kang H, Xu Q, et al. Sleep drives metabolite clearance from the adult brain. *Science* 2013;342:373–7.
35. Gangwisch JE, Heymsfield SB, Boden-Albala B, et al. Sleep duration as a risk factor for diabetes incidence in a large U.S. sample. *Sleep* 2007;30:1667–73.
36. Tasali E, Laproult R, Ehrmann DA, Van Cauter E. Slow-wave sleep and the risk of type 2 diabetes in humans. *PNAS* 2008;105:1044–9.
37. Bencza RM, Obermeyer WH, Thisted RA, Gillin JC. Sleep and psychiatric disorders: a meta-analysis. *Arch Gen Psychiatry* 1992;49:651–68.
38. Ohayon MM. Insomnia: a ticking clock for depression? *J Psychiatr Res* 2007;41:893–4.
39. Fung MM, Peters K, Redline S, et al. Decreased slow wave sleep increases risk of developing hypertension in elderly men. *Hypertension* 2011;58:596–603.
40. Collop NA, Anderson WM, Boehlecke B, et al. Clinical guidelines for the use of unattended portable monitors in the diagnosis of obstructive sleep apnea in adult patients. Portable Monitoring Task Force of the American Academy of Sleep Medicine. *J Clin Sleep Med* 2007;3:737–47.
41. Kirsch DB. PRO: sliding into home: portable sleep testing is effective for diagnosis of obstructive sleep apnea. *J Clin Sleep Med* 2013;9:5–7.