Sleep/wake Classification using Head Actigraphy, Snoring and Airflow Signals

Djordje Popovic[1], Vladislav Velimirovic[1], Indu Ayappa[2], Daniel Levendowski[1], Tim Zavora[1], David Rapoport[2], Philip Westbrook[1]

Advanced Brain Monitoring[1], New York University School of Medicine[2]

Introduction: Actigraphy is often used in ambulatory devices to differentiate wakefulness from sleep but its accuracy is limited by the inability to distinguish quiet wakefulness from sleep and active wakefulness from arousals resulting from sleep disordered breathing. This study investigates the accuracy of a novel algorithm which combines actigraphy, nasal flow and snoring obtained from the forehead in order to differentiate sleep from wake.

Method: One-hundred and nine subjects wore the ARES Unicorder (Advanced Brain Monitoring, Carlsbad, CA) concurrent during laboratory polysomnography (PSG). The algorithm was developed using a model development set of 25 subjects (RDI = 28 ± 26; range: 3-109) and cross-validated on 84 subjects (RDI = 28 ± 26; range: 1-103). PSG recordings were manually scored according to the AASM criteria. The sleep/wake algorithm begins with actigraphy to classify 30-second epochs as Wake based on the intensity and duration of head movements and/or the upright head position. Periods detected as wake or classified as wake subsequent to a prolonged period of gross movement are then reclassified to behavioral sleep when single-peak snores align with breaths in the airflow signal or when at least three consecutive airflow detected apneas/hypopneas are recognized in a sequence.

Epoch-by-epoch comparisons (n=59,652) were then performed and sensitivity, specificity, agreement, kappa, and positive predictive value (PPV) calculated for each subject's epoch, and then averaged across subjects. PSG-ARES differences in sleep latency, total sleep time and sleep efficiency were tested with paired t-test, and impact of RDI on algorithm performance with Pearson correlation coefficient and ANOVA.

Table 1. Comparison of behavior vs. electrophysiological detection of Sleep/Wake

	Cross Validation (n=84)			Model Development (n=25)		
	Mean + SD	Min	Max	Mean + SD	Min	Max
Sleep Onset – min. diff.	2.9 + 8.6	-13	47	6.7 + 13.7	-7	43
Total Sleep Time – min. diff.	-.80 + 36.3	-83	95	-0.5 + 34.1	-71	73
Sleep Efficiency - % diff	1 + 10%	-26%	21%	2 + 12%	-32%	16%
Sensitivity - Wake	.66 + .19	0.19	1.00	.70 + .19	0.36	0.98
Specificity - Sleep	.91 + .06	0.67	0.99	.91 + .07	0.72	0.99
Agreement	.85 + .06	0.69	0.97	.85 + .08	0.68	0.95
Kappa	.52 + .14	0.13	0.84	.58 + .15	0.28	0.82
PPV - actigraphy only	.64 + .21	0.10	0.97			
PPV - plus snoring & airflow	.67 + .19	0.21	0.97			

Results: The algorithm's performance was similar for the model development and cross validation data sets (Table 1). When reviewing the PSG scoring of wake we found that 3.6% of epochs (1.6% in subjects with RDI<20 and 6.8% in those with RDI>21) contained clear

obstructive respiratory events and/or snoring were scored as awake. There were no significant differences in total sleep time or sleep efficiency, however sleep onset was significantly underestimated. The addition of changes in snoring and airflow increased the PPV as compared to using actigraphy alone without impairing sensitivity (<1% change in all subjects). When the data were stratified by RDI severity (Table 2), there were no significant differences in the performance of the algorithm, although those with an RDI > 30 showed the greatest increase in positive predictive value (of wake) as a result of inclusion of airflow and snoring.

Table 2. Comparison of behavior vs. electrophysiological Sleep/Wake by RDI severity

	Mean + SD (n = 84)		
Model Development Data Set	RDI < 15	RDI 16-30	RDI >30
Sleep Onset - minutes diff	2.4 + 8.3	3.6 + 8.4	3.1 + 9.5
Total Sleep Time - Minutes diff	-2.3 + 36.4	8.2 + 45.5	-5.6 + 27.8
Sleep Efficiency - % diff	-1 + 9%	1 + 12%	-3 + 10%
Sensitivity - Wake	.66 + .19	.68 + .23	.65 + .16
Specificity - Sleep	.92 + .06	.89 + .08	.91 + .06
Agreement	.86 + .06	.84 + .07	.84 + .07
Kappa	.53 + .14	.49 + .15	.52 + .14
PPV actigraphy only	.70 + .16	.61 + .24	.59 + .23
PPV all signals	.70 + .16	.62 + .23	.67 + .20

The model development and cross-validation data sets were combined to present the overall sleep onset, total sleep time and sleep efficiency patterns (Figure 1). These data were stratified (< or > 2.5 hours) to test for differences in classification accuracy attributed to short PSG recording times resulting from split night studies. Short study times had a significant impact on the difference between the electrophysiological vs. behavior measures of total sleep time (-13 $\pm$ 23 SD minutes for short vs. 2.6 $\pm$ 38 for long, $p < 0.01$) and sleep efficiency (-1 $\pm$ 13% vs. 0 $\pm$ 9%, $p < 0.01$).

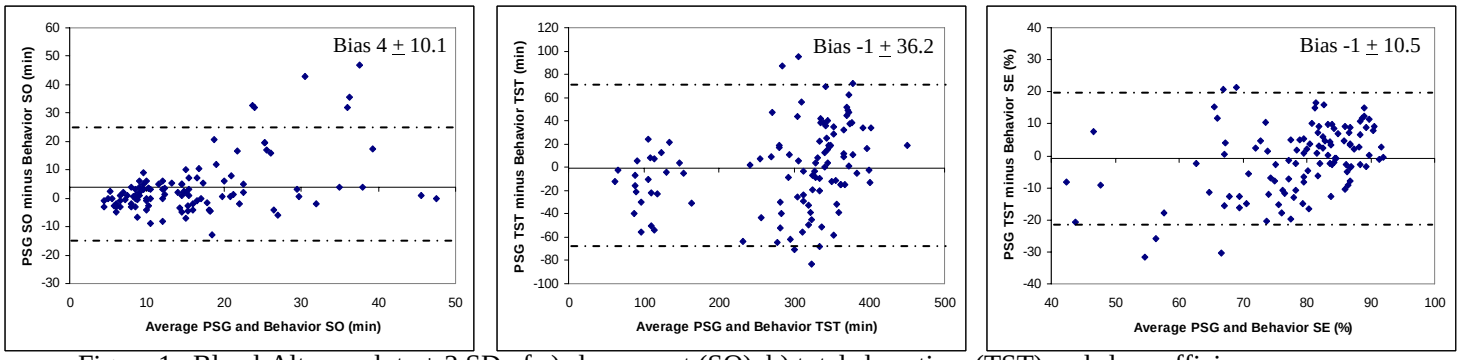

Figure 1. Bland-Altman plots $\pm$ 2 SD of a) sleep onset (SO), b) total sleep time (TST) and sleep efficiency (SE) (n=109).

Conclusion: Behavioral measures can be used to accurately differentiate wake from sleep. The accuracy of sleep/wake estimates are impacted, as compared to PSG, when split night studies are included in the analysis. Combining actigraphy with changes in airflow and snoring improves the detection of wake, especially in patients with severe RDI values.

Support: NIH SBIR R44HL068463-05 and 2R44-DE016772-0