

SCIENTIFIC INVESTIGATIONS

The Validity of Wrist Actimetry Assessment of Sleep With and Without Sleep Apnea

David Wang, Ph.D.^{1,2}; Keith K. Wong, MB. BS.^{1,2}; George C. Dungan II, RPSGT¹; Peter R. Buchanan, M.D.^{1,2}; Brendon J. Yee, Ph.D.^{1,2};
Ronald R. Grunstein, M.D., Ph.D.^{1,2}

¹Woolcock Institute of Medical Research, University of Sydney, Sydney, Australia; ²Department of Respiratory & Sleep Medicine, Royal Prince Alfred Hospital, Camperdown, Sydney, Australia

Introduction: It is unclear whether actimetry can be reliably used to measure sleep in severe obstructive sleep apnea (OSA) patients. We compared polysomnography (PSG) with actimetric assessment of sleep on an epoch-by-epoch basis in subjects with and without OSA.

Methods: 21 subjects were recorded with simultaneous overnight standard PSG and actimetry.

Results: 10 subjects with apnea-hypopnea index (AHI) <10 ($6.5 \pm 2.8/h$) were classified as non-OSA subjects and 11 subjects with AHI >10 ($42.0 \pm 27.3/h$) were classified as OSA subjects. Overall sensitivity and specificity for actimetry to identify sleep was 94.6% and 40.6%, respectively, with an overall mean sleep/wake simple agreement of 84.6% and kappa of 0.38. There was no difference in agreement between non-OSA and OSA subjects (simple agreement: 83% vs. 86%, $p = 0.73$; kappa: 0.35 vs. 0.40, $p = 0.73$). The kappa agreement did not correlate with PSG arousal index ($r = -0.21$, $p = 0.36$) but declined with reduced sleep efficiency ($r = 0.66$, $p = 0.001$). There was no systematic difference (all $p > 0.40$) between actimetry and PSG in sleep

latency, total sleep time and sleep efficiency, although correlations between the measurements using the two techniques were generally poor. However, while actimetry systematically underestimated wake after sleep onset (WASO) (35.5 ± 18.8 vs. 59.4 ± 35.1 , $p = 0.009$), fragmentation index measured by actimetry only underestimated arousal index measured by PSG in OSA patients (23.9 ± 17.8 vs. 33.1 ± 18.5 , $p = 0.04$).

Conclusions: Contrary to prior reports, epoch-by-epoch comparison of sleep/wake scoring showed similar fair agreement between actimetry and PSG in subjects with or without OSA. Fragmentation index by actimetry may underestimate arousals caused by respiratory events and offer misleading results in severe OSA patients.

Keywords: Actigraphy, PSG, validity, accuracy, agreement, sleep apnea, OSA, fragmentation index, arousal index

Citation: Wang D; Wong KK; Dungan II GC; Buchanan BR; Yee BJ; Grunstein RR. The validity of wrist actimetry assessment of sleep with and without sleep apnea. *J Clin Sleep Med* 2008;4(5):450-455.

There has been a substantial increase in the number of studies using actimetry to estimate sleep in recent years.¹ Compared to polysomnography (PSG), actimetry is easier to use, less invasive, and substantially less expensive. It has been commonly used as an alternative to PSG to measure sleep and evaluate effect of a treatment on sleep in clinical trials.² However, it is unclear whether actimetry can be reliably used to measure sleep in severe obstructive sleep apnea (OSA) patients.

In a recent AASM practice parameter, actimetry is cited as an alternative method to measure total sleep time in OSA patients when PSG is not available.¹ The estimation of sleep time is particularly useful in portable sleep apnea tests combined with a respiratory monitoring device.^{1,3,4} Despite this approach, actimetry has been suggested in various reviews and practice parameters to be less reliable in detecting disturbed sleep, and the validity of actimetry declines with increasing severity of OSA.^{1,2,4} This claim could significantly limit the utility of actimetry, as OSA is one of the most common sleep disorders

needing evaluation. The claim for reduced actimetry validity in sleep disordered breathing was based on a few studies using different actimetry techniques.^{5,6} Conversely, there were also studies reporting close agreement between PSG and actimetry in predominantly OSA patients, although no normal control group was used as a comparator.^{7,8} It has been suggested that differences in actimetry devices, data collection strategies and scoring algorithms may produce different results for the same activity.^{2,9} Given the continuing improvement of actimetry technology, we considered whether it was appropriate to extend previous validity claims to actimetry studies that have used different or newer technology.

In this study, we compared sleep parameters measured by PSG and a widely-used actimetry device on an epoch-by-epoch basis, in subjects with and without sleep apnea. The primary goal was to assess if the validity of actimetry decrease with an increase of OSA severity. In addition, we compared the fragmentation index generated by actimetry with the arousal index scored on the PSG with and without the presence of OSA, as such a comparison has not been described previously.¹

METHODS

Twenty-one subjects participated in the study. All provided informed consent. Ten subjects were heavy snorers with frequent

Submitted for publication February, 2008

Accepted for publication May, 2008

Address correspondence to: David Wang, Department of Respiratory & Sleep Medicine, Royal Prince Alfred Hospital, Camperdown, Sydney 2050, Australia; Tel: 0061-2-9515 5048; Fax: 0061-2-9515 7196; E-mail: dwang@med.usyd.edu.au

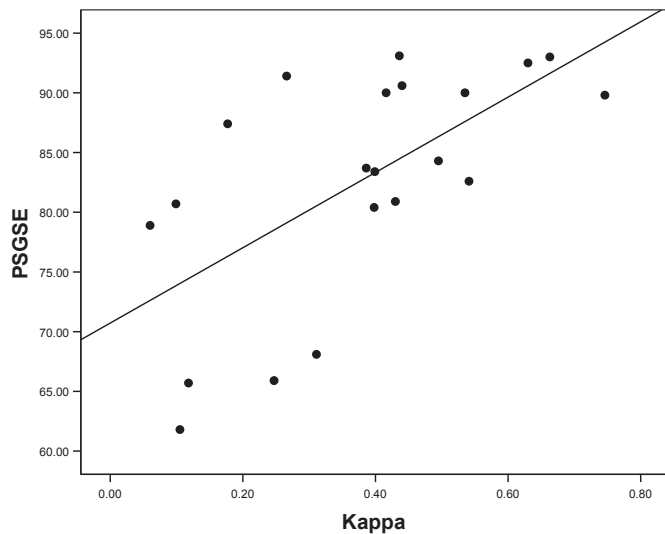


Figure 1—Correlation of sleep efficiency and accuracy of actimetry sleep/wake detection. Pearson correlation coefficient was tested and significant correlation was found ($r = 0.66$, $p = 0.001$). Possible interaction of OSA grouping in the correlation was tested. No significant interaction was found ($p = 0.76$). SE = Sleep efficiency.

reported apnea events during sleep, and 11 were healthy volunteers with no self-reported sleep disordered breathing symptoms. However, the eventual grouping of non-OSA or OSA was based on the AHI value rather than initial recruitment classification. Each subject underwent a full-night PSG study with simultaneous recording of actimetry (AW64, Mini-Mitter, Respironics, USA). The PSG and actimetry were time-synchronized before the study. Both PSG and actimetry data were coded and analyzed by an experienced and blinded sleep scientist (DW, RPSGT). The study was approved by the institutional research and ethics committee (University of Sydney HREC number 8609).

PSG

Overnight PSG was recorded using a standard technique, with measurements of electroencephalogram (EEG) (C_3-A_2 , C_4-A_1 , O_1-A_2 , O_2-A_1), bilateral electrooculogram (EOG), chin electromyogram (EMG), tibialis anterior EMG, electrocardiogram (ECG), nasal air pressure, percentage oxygen saturation (SpO_2), snoring, and body position. The PSG studies were performed using the Alice 5 system (Respironics, PA, USA). Sleep staging was scored with Rechtschaffen and Kales criteria.¹⁰ Respiratory events and arousals were scored according to standard Chicago¹¹ and ASDA criteria,¹² respectively. Apnea-hypopnea index (AHI) was calculated by dividing the total number of apneas and hypopneas by the total sleep time (hours).

Actimetry

Each subject wore an actimeter on the nondominant hand during sleep. Epoch length was set at 30 sec to match PSG setting. “Light off” and “Light on” time of the PSG recording were set as the start and the end of the “Rest Interval” of actimetry. The standard factory-default algorithm was used for the sleep interval detection. The parameters were: wake threshold was

Table 1—Age, BMI, and AHI of Subjects with and without OSA

	Age (yr)	BMI (kg/m ²)	AHI (/h)
Overall (n = 21)	38.9 ± 13.0	27.5 ± 6.1	25.1 ± 26.6
non-OSA (n = 10)	32.2 ± 12.7	24.6 ± 4.8	6.5 ± 2.8
OSA (n = 11)	45.0 ± 10.4	30.2 ± 6.2	42.0 ± 27.3

set as “medium”; sleep interval detection algorithm was set as “immobile minutes.” Both immobile minutes for sleep onset and end were set as 10 min.

The AW64 software (Actiware 5.0, Respironics, PA, USA) scores each epoch as either sleep or awake by comparing activity counts for the epoch in question and those immediately surrounding it (± 2 min) to a threshold value set by the researcher (“Medium”/Wake threshold value = 40). The Medium threshold algorithm has been validated in a previous technical report.¹³ If the number of counts exceeds the threshold, the epoch is scored as wake. If it falls below or equal to the threshold, the epoch is scored as sleep. Fragmentation Index is defined as “the sum of ‘percent mobile’ and ‘percent immobile bouts less than 1-min duration to the number of immobile bouts’, for the given interval.” Number of mobile bouts or immobile bouts is defined as “the total number of continuous blocks, one or more epochs in duration, with each epoch of each block scored as MOBILE or IMMOBILE (respectively), between the start time and the end time of the given interval.” More details are given in the instruction manual.¹⁴

An accurate synchronization between actimetry and PSG is critically important in validity studies.² In our study, we carefully synchronized the time on all the PSGs and actimetry recordings. In addition, when matching actimetry data with PSG data between the “light off” and “light on” time, we always tested the agreements a number of times within ± 2 min time range. The range of the agreements was always “bell” shaped. The peak agreement value was adopted in the final statistics.

Statistical Methods

Data are presented as mean \pm SD. Epoch-by-epoch sleep/wake agreement between actimetry and PSG was evaluated by percentage of epoch in agreement, sensitivity, specificity, and Cohen’s kappa statistic.¹⁵ Sensitivity is defined as the proportion of PSG scored “sleep” epochs that were rated as “sleep” with actimetry. Specificity is defined as the proportion of PSG scored “wake” epochs that were rated as “wake” with actimetry. The agreement values between subjects with and without OSA were compared using the Mann Whitney U test. Intraclass correlation coefficient (ICC) was used to evaluate accuracy of the sleep parameters evaluated by actimetry and PSG. ICCs were for average measures and tested with 2-way mixed effects model. We used a consistency definition of the ICC, i.e., the between-measure variance is excluded from the denominator variance.¹⁶ Sleep parameter comparisons between actimetry and PSG were performed using paired *t*-test. The Pearson correlation coefficient was used to assess association between two normally distributed continuous variables. Association between non-normally distributed variables was evaluated by Spearman rho. A *p* value < 0.05 was considered as statistically significant.

Table 2—Epoch-by-Epoch Sleep/Wake Agreement between Actimetry and PSG

	Agreement	Sensitivity	Specificity	kappa
Overall (n = 21)	0.85 ± 0.09	0.95 ± 0.05	0.41 ± 0.21	0.38 ± 0.19
Non-OSA (n = 10)	0.83 ± 0.12	0.94 ± 0.07	0.41 ± 0.25	0.35 ± 0.24
OSA (n = 11)	0.86 ± 0.06	0.95 ± 0.04	0.40 ± 0.17	0.40 ± 0.15
p	0.73	0.73	0.89	0.73

Table 3—Comparison of PSG and Actimetry in Sleep Parameters

		Overall (n = 21)	Non-OSA (n = 10)	OSA (n = 11)
Sleep Latency (min)	PSG-Lat	24.74 ± 29.91	39.60 ± 37.58	11.23 ± 9.79
	Act-Lat	27.60 ± 53.77	41.70 ± 75.20	14.77 ± 17.70
	p	0.81	0.94	0.53
	ICC	0.35	0.20	0.31
Total Sleep Time (min)	PSG-TST	400.83 ± 52.83	391.65 ± 61.98	409.18 ± 44.31
	Act-TST	407.90 ± 62.33	392.75 ± 72.30	421.68 ± 51.24
	p	0.67	0.97	0.49
	ICC	0.27	0.06	0.44
Sleep Efficiency (%)	PSG-SE	82.58 ± 9.69	80.99 ± 11.80	84.03 ± 7.60
	Act-SE	84.41 ± 13.16	81.84 ± 16.34	86.74 ± 9.67
	p	0.59	0.89	0.47
	ICC	0.23	0.21	0.13
Arousal/Fragmentation Index (/h)	PSG-AI	24.40 ± 16.22	14.78 ± 3.03	33.15 ± 18.48
	Act-Frag	21.86 ± 15.04	19.58 ± 11.88	23.94 ± 7.77
	p	0.41	0.20	0.04*
	ICC	0.76	0.33	0.85
WASO (min)	PSG-WASO	59.4 ± 35.07	51.65 ± 36.83	66.45 ± 33.53
	Act-WASO	35.45 ± 18.84	30.1 ± 11.87	40.32 ± 22.99
	p	0.009*	0.13	0.04*
	ICC	0.18	0.23	0.3

p values were tested by paired *t*-test; ICC = intraclass correlation coefficient; ICCs were for average measures and tested with 2-way mixed effects model. We selected Type C ICC using a consistency definition; Lat = sleep onset latency; TST = total sleep time; SE = sleep efficiency; AI = arousal index; Act = actimetry; Frag = fragmentation index; WASO = wake after sleep onset.

RESULTS

Subjects' age, body mass index (BMI), and AHI are listed in Table 1. Ten subjects with AHI <10 were classified as non-OSA, and 11 subjects with AHI >10 were classified as OSA patients. One of the self-reported healthy volunteers had an AHI >10 and was therefore classified as an OSA subject. In detail, 3 subjects had AHI <5, and 7 subjects had AHI between 5 and 10. Four subjects had AHI between 10 and 30, and 7 subjects had severe OSA with an AHI >30.

The overall sensitivity and specificity for actimetry to identify sleep/wake were 94.6% and 40.6% respectively, with an overall mean simple sleep/wake agreement of 84.6% and kappa of 0.38. (Table 2) There was no difference in any actimetry validity measurement between non-OSA and OSA subjects. (Table 2) No correlation was found between AHI and kappa ($\rho = -0.095$, $p = 0.68$). Similarly, the kappa agreement did not correlate to PSG arousal index ($r = -0.21$, $p = 0.36$). However, the actimetry validity significantly declined with the decrease of sleep efficiency ($r = 0.66$, $p = 0.001$). (Figure 1)

With the exception of the arousal index/fragmentation index and wake after sleep onset (WASO), no significant differences were found between sleep measures derived from PSG and actimetry. (Table 3) This persisted for the whole group, as well as for the OSA and non-OSA subgroups. However, the low ICCs may suggest generally poor correlations between the 2 measurements

of sleep parameters regardless of sleep apnea severity. (Table 3; Figure 2) The differences between measurement methodologies are shown to be evenly distributed around the Bland Altman plot for sleep latency, total sleep time, and sleep efficiency (Figure 2), showing no obvious systematic differences between the methodologies, while still highlighting substantial differences for certain subjects in the study. Actimetry was found to systematically underestimate WASO compared to PSG with an overall mean difference of 24 minutes ($p = 0.009$). The mean underestimation is 26 minutes in OSA patients ($p = 0.04$) and 21 minutes in non-OSA subjects ($p = 0.13$). (Figure 2; Table 3)

Although the fragmentation index measured by actimetry tends to be higher than arousal index measured by PSG in non-OSA subjects (19.6 ± 11.9 actimetry vs. 14.8 ± 3.0 PSG, $p = 0.20$), the fragmentation index is significantly lower than PSG arousal index in OSA subjects (23.9 ± 17.8 actimetry vs. 33.1 ± 18.5 PSG, $p = 0.04$) (Table 3). The relationship is shown in Figure 3. Both arousal index and fragmentation index are significantly correlated with AHI. However, the relationship between the arousal index plot and fragmentation index plot is not parallel, i.e., one measurement is not systematically higher than the other. Instead, they demonstrate a crossover relationship. Graph (A) in Figure 3 shows that the intercept point is around an AHI value of 17. However, 2 clear outliers (marked with arrows) may bias the relationship. After they were discarded, the intercept point shifted to around AHI = 10 in Graph (B). (Figure 3)

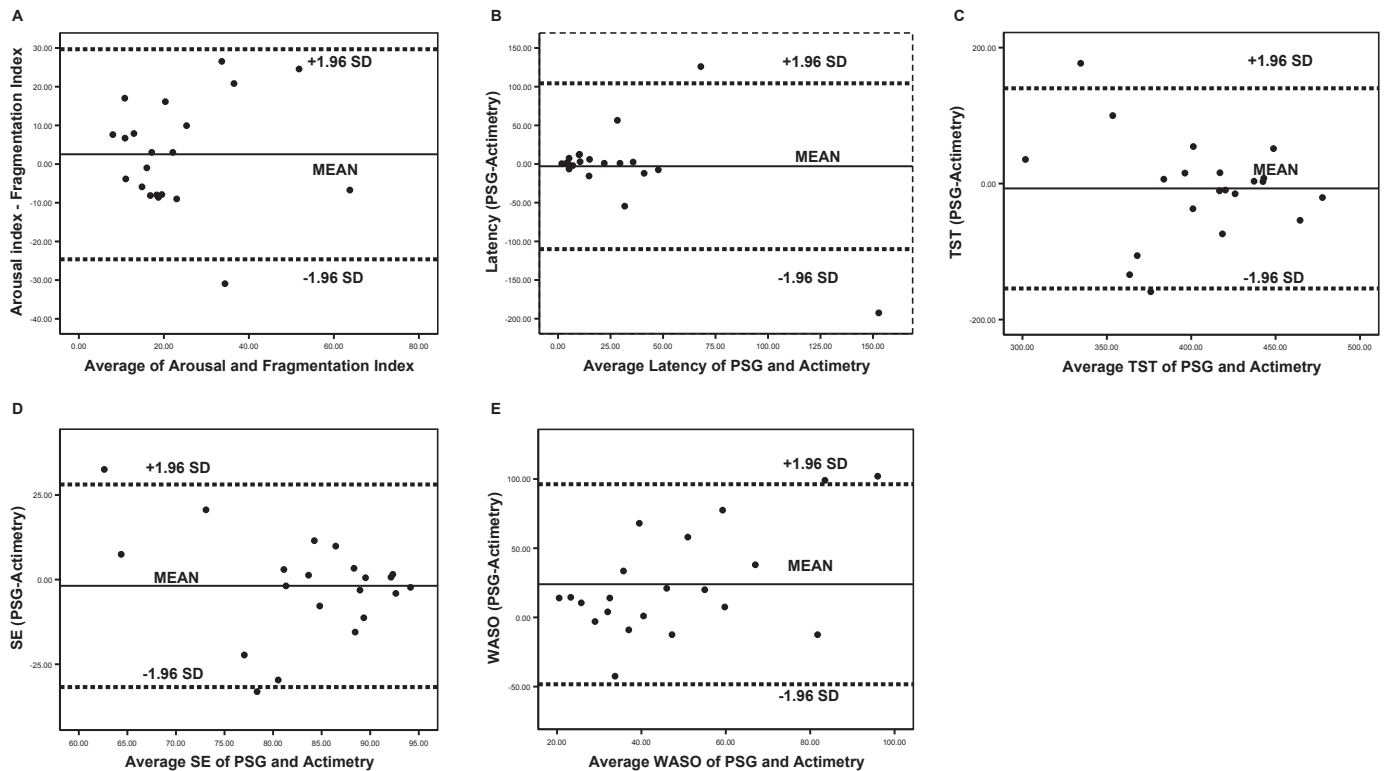


Figure 2—Bland Altman plot for sleep parameters between PSG and actimetry data. Latency = sleep onset latency, TST = total sleep time, SE = sleep efficiency, WASO = wake after sleep onset.

DISCUSSION

Unlike previous reports, we found that OSA severity and related respiratory arousal did not affect the accuracy of actimetry in detecting sleep/wake.^{1,2,5,6} We have found that the validity of actimetry declined linearly with the decrease of sleep efficiency. We also observed that fragmentation index recorded by actimetry underestimates PSG arousal index in OSA subjects but not in non-OSA subjects.

It has been concluded in the 2003 and 2007 AASM practice parameters^{1,4} and the 2003 AASM review paper² that actimetry is reliable and valid for detecting sleep in normal, healthy populations, but less reliable for detecting sleep in disturbed sleep or severe sleep apnea patients. Our non-OSA and OSA groups had an average AHI of 6.5/h and 42/h respectively. With such a large difference in sleep apnea severity, actimetry in the 2 groups still showed no systematic difference in detecting sleep/wake. Similarly, the arousal index did not correlate with the kappa agreement. We suspect these novel findings may be due, in part, to different actimetry technology/algorithm/analyses employed in prior versus the current study. The disparate findings may suggest that it is difficult to generalize validity results between various technologies and scoring algorithms, especially within the context of improving technology.

For example, a study in 2004 reported that the agreement of sleep-wake measured by actimetry as compared to PSG declined with the increase of OSA severity.⁵ The agreement ranged from 86% in the normal subject to 86%, 84%, and 80% in the patients with mild, moderate and severe OSA respectively. The difference is quite modest when compared to the difference in AHI. The actimeter used in that study was the Watch_PAT100 system

and the data were analyzed with ASWA software.⁵ In contrast, a recent study tested 181 adolescents with both PSG and actimetry.⁶ Seventeen with AHI >5 were classified into a sleep disordered breathing (SDB) group. However, even this group only had an average AHI of 9/h. The No-SDB group consisted of 164 subjects with AHI <5. The TST median ICC compared with PSG in No-SDB group was 0.55, 0.42, and 0.43 respectively for 3 different actimetry data modes; while the median ICC in SDB group was 0.00, 0.00, and 0.02 respectively. The actimetry data were collected using the Octagonal Sleep Watch 2.01 and analyzed using Action-W analysis software. Considering the small difference in AHI between the 2 groups, the difference in agreement between actimetry and PSG is substantial.⁶

There were 2 studies reporting close agreement between actimetry and PSG in predominantly OSA patients. In one study, the Pearson correlation coefficient between TST tested by actimetry (Cambridge Neurotechnology Ltd) and PSG was 0.90, $p < 0.0001$.⁷ The mean TST difference was 2.5 min. Subjects consisted of 24 patients with average AHI of 38.4/h. No epoch-by-epoch agreement was tested, and no normal control group was compared.⁷ In the other study, epoch-by-epoch comparison of sleep/wake by actimetry (AW4, Mini-Mitter Co.) and PSG showed an overall sensitivity of 0.96 and specificity of 0.38, which is similar to our results.⁸ The subjects consisted of 100 patients including 85 with SDB and 15 with other sleep disorders, including insomnia, narcolepsy, hypersomnia, and restless legs syndrome.⁸ The 15 non-SDB patients had a sleep/wake sensitivity of 0.96 and specificity of 0.42.⁸ From these 2 studies, we could not draw a definite conclusion about the validity of actimetry in OSA patients as none of the studies has a comparison normal control group.

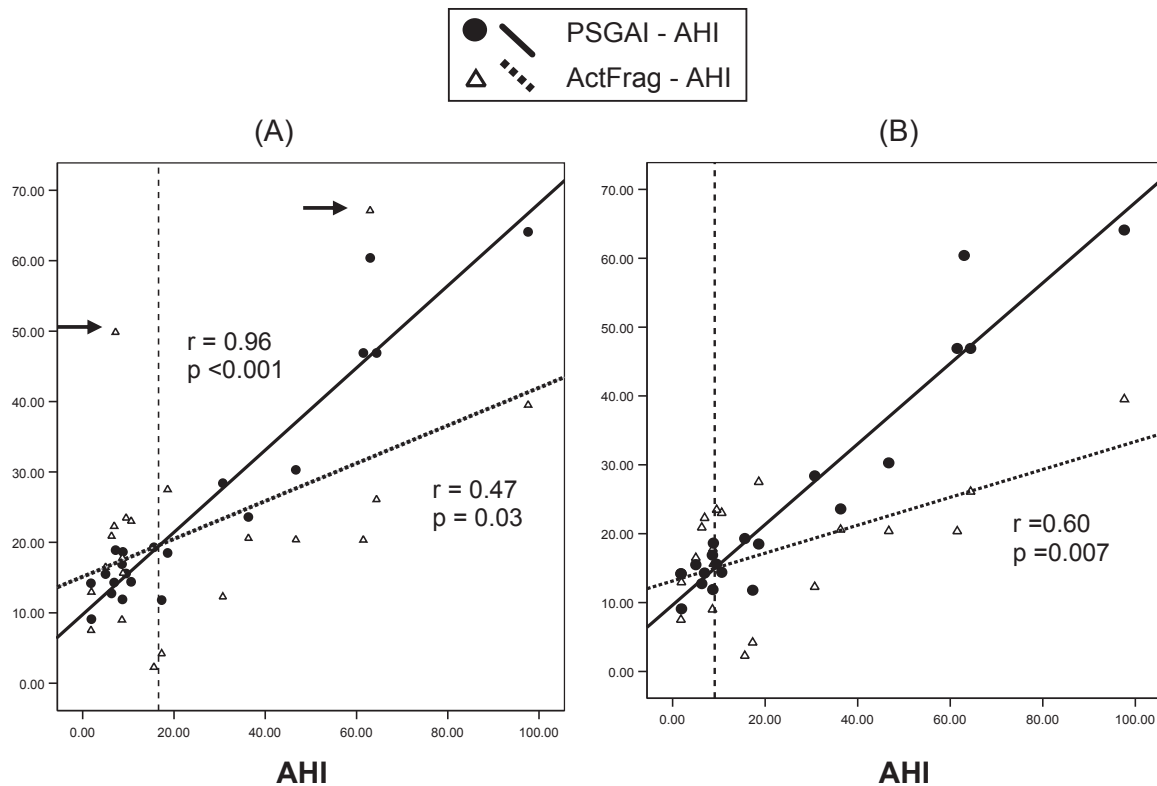


Figure 3—PSG Arousal Index and Actimetry Fragmentation Index plot against AHI. All the associations are significant. The relationship between the two plots is crossover and not parallel. Graph (A) shows the raw data with 2 outliers marked by arrows. The intercept point is around AHI = 17. Graph (B) shows the modified plot with the 2 outliers excluded. The intercept point shifted to around AHI = 10. AI = arousal index, ActFrag = actimetry-measured fragmentation index.

As suggested in the AASM review, we found a relationship between the decline of the accuracy of actimetry to detect sleep/wake and the decrease of sleep efficiency.² This may suggest that although the detection of sleep/wake in the technology used in our study is not sensitive to sleep apnea or respiratory arousal, it is sensitive to overall sleep efficiency reduction.

Another interesting finding from this study is the comparison of actimetry fragmentation index with PSG arousal index, which has not been well described before.¹ In this study, the measurement of sleep/wake, sleep latency, TST and sleep efficiency in actimetry is not systematically different to PSG, although their correlation is generally poor regardless of the presence or absence of OSA. However, Figure 3 shows that fragmentation index in OSA subjects (AHI >10) underestimates arousal severity compared to PSG, and the bigger AHI, the larger the underestimation. Whereas in non-OSA subjects with AHI <10, fragmentation index tends to mildly overestimate arousals. The intercept of the crossover relationship (AHI = 10) coincides with a commonly used cut-off point for OSA (Figure 3). This cross-over relationship contrasts to common speculations that one measurement could be systematically higher or lower than the other measurement (parallel relationship). Those results may suggest that actimetry is not fit to be singly used as a diagnostic tool for OSA given the significant underestimations of respiratory arousals in severe OSA patients. The systematic underestimation of WASO in actimetry may further strengthen this point.

A limitation of the study is the relatively small sample size although we have reached statistical significance in some im-

portant parameters. It is possible that with an increased sample size, the validity of actimetry might show a decrease with the increase of the sleep apnea severity. However, we suspect that even if this were to occur, the magnitude of the effect would be weak. In contrast, the effect from a reduced sleep efficiency is quite robust as we have demonstrated in our study. Further study to confirm our findings with a larger sample size and a broader spectrum of OSA severity is certainly appropriate. Another limitation of the study is the actimetry algorithm used. As the focus of the study was the difference of actimetry validity with or without OSA rather than purely as a technical validity assessment, we only applied the most commonly used setting—medium threshold (the factory default). However the measurement differences using low-, medium- or high-threshold algorithms were minor as described in previous studies.⁸ Another practical issue is that it is often difficult to obtain an accurate “lights out” time for actimetry analysis outside the more readily controlled environment of a sleep laboratory. “Percent time asleep” (from the first sleep onset to the last wake) is therefore commonly used in actimetry studies. While “percent time asleep” is similar to “sleep efficiency” in nature, we did not report “percent time asleep” separately.

A statistical paradox is shown in the sleep/wake agreement between actimetry and PSG, similar to what Feinstein and Cicchetti termed “the first paradox of kappa”¹⁷ (Table 2). The 2 measurements have a good percentage agreement (83% to 86%) but only fair kappa agreement (0.35-0.40). The kappa coefficient corrects for the amount of agreement that might be expect to occur by chance alone. Owing to a high proportion

of epochs of sleep (about 80%), this correction for chance may lead to a low value of kappa even if the 2 measurements appear to have high agreement.¹⁷

In conclusion, in contrast to previous reports, our data showed that the validity of actimetry is not sensitive to the severity of sleep apnea, but is sensitive to overall reduction of sleep efficiency. Fragmentation index by actimetry may underestimate arousals caused by respiratory events and give misleading results in severe sleep apnea patients.

ACKNOWLEDGMENT

The study was supported by NHMRC CCRE in Respiratory and Sleep Medicine; RACP CONROD Fellowship (Dr. Keith Wong); NHMRC Practitioner Fellowship (Prof. Ronald Grunstein).

Financial Support: NHMRC CCRE in Respiratory and Sleep Medicine; RACP CONROD Fellowship (Dr. Keith Wong); NHMRC Practitioner Fellowship (Prof. Ronald Grunstein).

DISCLOSURE STATEMENT

This was not an industry supported study. The authors have indicated no financial conflicts of interest.

REFERENCES

- Morgenthaler T, Alessi C, Friedman L, et al. Practice parameters for the use of actigraphy in the assessment of sleep and sleep disorders: an update for 2007. *Sleep* 2007;30:519-29.
- Ancoli-Israel S, Cole R, Alessi C, Chambers M, Moorcroft W, Pollak CP. The role of actigraphy in the study of sleep and circadian rhythms. *Sleep* 2003;26:342-92.
- Standards of Practice Committee. Practice parameters for the use of actigraphy in the clinical assessment of sleep disorders. American Sleep Disorders Association. *Sleep* 1995;18:285-7.
- Littner M, Hirshkowitz M, Kramer M, et al. Practice parameters for using polysomnography to evaluate insomnia: an update. *Sleep* 2003;26:754-60.
- Hedner J, Pillar G, Pittman SD, Zou D, Grote L, White DP. A novel adaptive wrist actigraphy algorithm for sleep-wake assessment in sleep apnea patients. *Sleep* 2004;27:1560-6.
- Johnson NL, Kirchner HL, Rosen CL, et al. Sleep estimation using wrist actigraphy in adolescents with and without sleep disordered breathing: a comparison of three data modes. *Sleep* 2007;30:899-905.
- Gagnadoux F, Nguyen XL, Rakotonanahary D, Vidal S, Fleury B. Wrist-actigraphic estimation of sleep time under nCPAP treatment in sleep apnoea patients. *Eur Respir J* 2004;23:891-5.
- Kushida CA, Chang A, Gadkary C, Guilleminault C, Carrillo O, Dement WC. Comparison of actigraphic, polysomnographic, and subjective assessment of sleep parameters in sleep-disordered patients. *Sleep Med* 2001;2:389-96.
- Gorny SW, Spiro JR. Comparing different methodologies used in wrist actigraphy. *Sleep Review* 2001; Summer:40-42.
- Rechtschaffen A, Kales A. A manual of standardized terminology, techniques and scoring systems for sleep stages of human subjects. Washington, D. C.: Public Health Services, U.S. Government Printing Office, 1968.
- AASM Task Force. Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement technique in clinical research. *Sleep* 1999;22:667-89.
- American Sleep Disorders Association. 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.
- Oakley NR. Validation with polysomnography of the sleepwatch sleep/wake scoring algorithm used by the actiwatch activity monitoring system. Technical Report to Mini Mitter Co., Inc., 1997.
- Respironics. Actiware Software: Actiwatch Instruction Manual, software version 5.0. Oregon, USA: Mini Mitter Company, Inc. A Respironics, Inc. Company, 2005.
- Cohen J. Coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960;20:37-46.
- McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods* 1996;1:30-46.
- Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. *J Clin Epidemiol* 1990;43:543-9.