JCSM
Journal of Clinical Sleep Medicine

SCIENTIFIC INVESTIGATIONS

Comparison of Commercial Wrist-Based and Smartphone Accelerometers, Actigraphy, and PSG in a Clinical Cohort of Children and Adolescents

Elicia Toon, BA(Hons)[1,2]; Margot J. Davey, MBBS[1,3]; Samantha L. Hollis, BA(Hons)[1]; Gillian M. Nixon, MD[1,3]; Rosemary S.C. Horne, PhD[1]; Sarah N. Biggs, PhD[1]

[1]The Ritchie Centre, Hudson Institute of Medical Research and Department of Paediatrics, Monash University, Melbourne, Australia; [2]School of Psychological Sciences, Monash University, Melbourne, Australia; [3]Melbourne Children's Sleep Centre, Monash Children's Hospital, Melbourne, Australia

Study Objectives: To compare two commercial sleep devices, an accelerometer worn as a wristband (UP by Jawbone) and a smartphone application (MotionX 24/7), against polysomnography (PSG) and actigraphy (Actiwatch2) in a clinical pediatric sample.

Methods: Children and adolescents (n = 78, 65% male, mean age 8.4 ± 4.0 y) with suspected sleep disordered breathing (SDB), simultaneously wore an actiwatch, a commercial wrist-based device and had a smartphone with a sleep application activated placed near their right shoulder, during their diagnostic PSG. Outcome variables were sleep onset latency (SOL), total sleep time (TST), wake after sleep onset (WASO), and sleep efficiency (SE). Paired comparisons were made between PSG, actigraphy, UP, and MotionX 24/7. Epoch-by-epoch comparisons determined sensitivity, specificity, and accuracy between PSG, actigraphy, and UP. Bland-Altman plots determined level of agreement. Differences in bias between SDB severity and developmental age were assessed.

Results: No differences in mean TST, WASO, or SE between PSG and actigraphy or PSG and UP were found. Actigraphy overestimated SOL (21 min). MotionX 24/7 underestimated SOL (12 min) and WASO (63 min), and overestimated TST (106 min) and SE (17%). UP showed good sensitivity (0.92) and accuracy (0.86) but poor specificity (0.66) when compared to PSG. Bland-Altman plots showed similar levels of bias in both actigraphy and UP. Bias did not differ by SDB severity, however was affected by age.

Conclusions: When compared to PSG, UP was analogous to Actiwatch2 and may have some clinical utility in children with sleep disordered breathing. MotionX 24/7 did not accurately reflect sleep or wake and should be used with caution.

Keywords: actigraphy, accelerometer, polysomnography, pediatrics

Citation: Toon E, Davey MJ, Hollis SL, Nixon GM, Horne RS, Biggs SN. Comparison of commercial wrist-based and smartphone accelerometers, actigraphy, and PSG in a clinical cohort of children and adolescents. *J Clin Sleep Med* 2016;12(3):343–350.

INTRODUCTION

Short sleep duration or sleep disruption during childhood, either due to a physiological sleep disorder or psychosocial basis, have adverse effects on brain development, cognitive performance, behavioral functioning, and psychological well-being.[1–7] Appropriate treatment of sleep problems in children and adolescents depends on accurate assessment. The current gold standard for assessing sleep and diagnosing sleep disorders in children is polysomnography (PSG).[8,9] This technique, which uses sophisticated technology to assess brain, cardiovascular and respiratory activity during sleep usually in a laboratory environment, is expensive and labor-intensive and typically only provides one or two nights of information, which may not be reflective of sleep in the home.[10]

Actigraphy, which use accelerometer technology to provide estimates of sleep and wake based on the level of activity, is used to provide objective assessments of sleep-wake patterns over long periods of time[11,12] and assist in diagnosis and treatment monitoring of sleep disorders such as delayed sleep phase syndrome and behavioral insomnia.[13] Validation studies in infants, children and adolescents have shown that actigraphy is sensitive in assessing actual sleep, however over- or underestimates wake after sleep onset, providing a poor estimate of sleep disruption when compared to PSG.[14–17] While actigraphy is less expensive, labor-intensive, and more reflective of natural sleep patterns than PSG, it still requires specialized software and expertise to analyze and interpret.[16]

A number of commercial devices that claim to assess sleep duration and quality are now available. Some use accelerometer technology similar to actigraphy, such as the Fitbit Ultra (Fitbit Inc. San Francisco, CA) and UP manufactured by Jawbone (Jawbone, San Francisco, CA), which are small devices, typically worn around the wrist. Others come in the form of

BRIEF SUMMARY

Current Knowledge/Study Rationale: Appropriate treatment of sleep problems in children and adolescents depends on accurate assessment. Inexpensive and publicly available commercial devices that claim to measure sleep duration and quality have risen in popularity, but are yet to be adequately assessed.

Study Impact: Commercial wrist-based accelerometer devices carry the same biases as actigraphy, but could be similarly used to assess sleep-wake patterns in children. Accelerometer smartphone applications are not as sensitive as wrist-based devices and should be considered carefully before used in clinical settings.

smartphone applications such as the MotionX 24/7 (Fullpower Technologies, Inc. Santa Cruz, CA), Sleep Cycle (Northcube, Göteborg, Sweden), and SleepBot (Sleepbot, New York, NY). The public availability, low cost and easily interpreted software make these devices an attractive alternative to actigraphy; however, validation against PSG is scarce in children,[18] adolescents,[19] and adults.[20]

The aim of this study was to compare two widely available commercial sleep devices—an accelerometer worn as a wristband (UP) and a smartphone application (MotionX 24/7) against PSG and actigraphy (Actiwatch2).

METHODS

Participants

Participants, aged 3 to 18 years, were recruited from a clinical cohort with suspected obstructive sleep apnea (OSA) scheduled for overnight PSG. Children with conditions that affect motor control and limb movement were not recruited. The Monash Health and Monash University Human Research Ethics Committees granted ethical approval. Written informed consent from parents/guardians and informed verbal assent from participants was obtained before the study. No monetary incentive was provided.

Protocol

Overnight PSG was performed in the Melbourne Children's Sleep Centre using Compumedics Sleep Systems (Compumedics, Melbourne, Australia). Electroencephalography (F3-M2, F4-M1, C3-M2, C4-M1, O1-M2, O2-M1), left and right electroculograms, submental and leg electromyograms, and electrocardiogram were recorded. Nasal pressure and oronasal airflow (nasal cannula and oronasal thermistor, Compumedics, Melbourne, Australia), thoracic and abdominal breathing movements (Resp-ez Piezo-electric sensor, EPM systems, Midlothian, VA, USA), blood oxygen saturation (Biox 3700e Pulse Oximeter, Ohmeda, Louisville, CO, USA), and transcutaneous carbon dioxide (TINA, TCM3, Radiometer, Denmark, Copenhagen) were recorded to identify respiratory events.

An actigraph (Actiwatch2, Philips Respironics, Pittsburgh, USA) and a UP (first release original version) were placed side-by-side on the non-dominant wrist. The MotionX 24/7 smartphone application installed onto an iPhone 4 (Apple Inc., Cupertino, CA) was set to start recording sleep data and placed under the bed sheet near the shoulder as recommended by the application developer. The UP and MotionX 24/7 were activated at lights out and deactivated at lights on.

As the study is a within-subjects design, the time clocks of actigraphy and the iPhone were synchronised to the PSG acquisition computer and recorded simultaneously with PSG over the sleep period. Technical specifications could not be obtained regarding the algorithms used by UP and MotionX 24/7 to calculate sleep and wake and thus are not reported.

Data Analyses

PSG studies were manually sleep-staged into 30-s epochs and respiratory events scored by experienced pediatric sleep technologists according to standardized rules.[21,22] The obstructive respiratory disturbance index (ORDI) was defined as the total number of obstructive apneas, mixed apneas, obstructive hypopneas, and respiratory event related arousals/h of total sleep time (TST). The arousal index was defined as the total number of cortical spontaneous, respiratory, and periodic leg movement arousals/h of TST.

Actigraphic data were downloaded using Actiware software (version 6.0.2) analyzed at a medium wake threshold setting (40 activity counts per epoch), shown to provide the best balance between sensitivity (ability to detect true sleep), specificity (ability to detect true wake), and accuracy (ability to detect both sleep and wake).[11] Epochs were coded as sleep when the total activity count per epoch was below the wake threshold level. UP was re-synchronized to the iPhone for each subject to eliminate any confounding effects of continual recording. At the completion of the sleep period, data were downloaded from the wrist device to the UP application on the iPhone.

The minimum epoch length for UP is one minute, therefore one-minute epoch-by-epoch matching was done for PSG, actigraphy and UP. As the wrist based devices were activated and de-activated earlier and later than PSG recording, respectively, the time of the first and last epoch on the PSG were applied to actigraphy and UP as the start and end times of the sleep period (sleep period time, SPT). For every minute, data for PSG, actigraphy and UP were manually coded as either sleep or wake (0 = wake, 1 = sleep). PSG provides data in 30-s epochs so each minute of PSG data was manually rescored as wake if either one or both of the 30-s epochs in each 1-min block were scored as wake, as per previous comparison studies.[15,17,23]

Epoch-by-epoch matching between PSG and MotionX 24/7 was not conducted as reducing PSG data to 5-min epochs, as is recorded on MotionX 24/7, is unreliable. To do this would require unacceptable subjective judgement regarding whether a 5-min recording, containing both sleep and wake, is coded as sleep or wake. Trimming the data to match with the start and end times of the PSG was also not possible for this device as only summary data are provided by the program.

Sleep outcome variables of sleep onset latency (SOL: number of minutes from start of the SPT to the first epoch of sleep); total sleep time (TST: number of minutes scored as sleep between start and end of SPT), wake after sleep onset (WASO: number of minutes scored as wake between start and end of SPT), and SE (TST/number of minutes within the SPT, expressed as a percentage) were analyzed.

Statistical Analyses

All statistical analyses were conducted using SPSS version 22.0 (SPSS, Inc., Chicago, IL). PSG data were ineligible for analysis for 4 participants (5%) as epoch-by-epoch sleep staging was not conducted due to excessive artifact. Sleep data were unavailable for 4% (n = 3) of actigraphy, 13% (n = 11) of UP, and 1% (n = 1) of MotionX 24/7 and were excluded from pairwise analyses. Missing data were due to participant influence (e.g., child taking off the UP during the night) or device malfunctions (e.g., actigraphy recording ceased due to battery malfunction).

Table 1—Calculation formula for sensitivity, specificity, and accuracy.

	PSG		
Actiwatch2/UP	**Sleep**	**Wake**	**Total**
Sleep	True Sleep (TS)	False Sleep (FS)	TS + FS
Wake	False Wake (FW)	True Wake (TW)	FW + TW
Total	TS + FW	FS + TW	TS + FS + TW + FW

Sensitivity = TS / (TS + FW); Specificity = TW / (FS + TW); Accuracy = (TS + TW) / (TS + FS + TW + FW).

Table 2—Cohort characteristics (n = 78, 65% male).

	Mean	Standard Deviation	Minimum	Maximum
Age (years)	8.4	4.0	3.1	17.1
BMI-Z score	0.9	1.2	−2.6	4.3
ORDI (events/h)	5.3	8.0	0.0	38.0
SpO_2 nadir (%)	91.5	4.7	72.0	97.0
Arousal Index (events/h)	14.4	5.8	4.0	35.1
PLMI (events/h)	1.6	3.2	0.0	20.0

BMI, body mass index; ORDI, obstructive respiratory disturbance index; SpO_2 nadir, lowest oxygen saturation level; PLMI, periodic limb movement index.

All data were checked for normality. SOL and WASO were positively skewed and were corrected using a logarithmic transformation.[24] Analyses were conducted on transformed data, however raw means are reported. Paired comparisons of SOL, TST, WASO, and SE between PSG, actigraphy, and UP were conducted using a paired-samples Student t-test. Paired comparisons between PSG and MotionX 24/7 were conducted using Wilcoxon signed-rank tests to account for the inability to match MotionX 24/7 epochs to PSG. Paired comparisons were only conducted for subjects with complete data for all 4 devices (n = 64). Multiple comparisons were controlled using Bonferroni adjustment with significance accepted at $p < 0.003$.

Epoch-by-epoch comparisons were conducted to determine the sensitivity, specificity, and accuracy of actigraphy and UP against PSG (**Table 1**). Epoch-by-epoch comparisons were conducted on pairwise matching.

Bland-Altman plots were used to examine the degree of agreement between PSG and actigraphy, and PSG and UP, for SOL, TST, WASO, and SE.[25] The mean difference (or bias) and lower and upper limits of agreement (95% confidence interval = mean difference ± 2SD)[25] are shown. A positive mean difference indicates an underestimation of the sleep parameter, while a negative difference indicates an overestimation. As a secondary analyses, ANCOVAs, with mean difference in sleep outcomes between PSG and UP as the dependent variable, were conducted to determine whether a difference in bias existed with increasing severity of SDB or increasing age. Children with SDB were categorized into primary snoring (ORDI ≤ 1 event/h), mild OSA (ORDI > 1–5 events/h), and moderate-severe OSA (ORDI > 5 events/h). Age was categorized into preschool (3.1–5.9 years), primary school (6.0–11.9 years), and adolescents (12.0–17.1 years).

Clinically acceptable agreement between PSG and actigraphy and PSG and UP were determined using standards based on previous research[15,26] and were defined as ≤ 30-min difference between devices for TST and WASO, and ≤ 5% difference for SE.

RESULTS

Cohort Characteristics

Mean age, BMI-Z score and respiratory parameters as measured on the PSG are shown in **Table 2**. Forty-nine percent of subjects had no comorbid conditions, 13% had chronic inflammatory diseases (e.g., asthma, eczema, recurrent tonsillitis), 13% had diagnosed behavioral disorders (e.g., autism spectrum disorder, anxiety, attention deficit hyperactivity disorder), and 26% had other comorbidities (e.g., Prader-Willi, Hunter Syndrome, cleft palate, Pierre Robin). Seventy-one percent were medication free at the time of the study, 18% were taking medications for asthma (e.g. fluticasone, salbutamol), and 11% were on other medications (e.g., methylphenidate, oxybutynin). Following the PSG, 31% were diagnosed with primary snoring (obstructive respiratory disturbance index: ORDI ≤ 1 event/h), 41% with mild OSA (ORDI > 1 and ≤ 5 events/h), and 28% with moderate to severe OSA (ORDI > 5 events/h). Six percent had a periodic limb movement index (PLMI) > 5 events/h.

Paired Comparisons

Results of the paired comparisons between PSG and actigraphy and PSG and UP are presented in **Table 3**. There were no significant mean differences between PSG and UP on any global sleep measure. Actigraphy significantly underestimated SOL by an average of 21 minutes compared to PSG.

SOL was significantly longer (21 min) and WASO was significantly shorter (13 min) when measured by UP compared to actigraphy. TST and SE showed no difference between these 2 devices.

Results of the paired comparisons between PSG and MotionX 24/7 are displayed in **Table 4**. Even when accounting for

Table 3—Results of parametric paired comparisons of sleep characteristics between PSG and actigraphy, PSG and UP, and UP and actigraphy.

	Devices		
Sleep Characteristics	**PSG**	**UP**	**Actiwatch2**
SOL (min)	34 (29)	34 (30)	13 (16)[a*,b*]
TST (min)	395 (76)	404 (70)	412 (45)
WASO (min)	82 (64)	73 (60)	86 (41)[b*]
SE (%)	77 (14)	79 (13)	81 (8)

All data are expressed as mean (± standard deviation). Mean sleep period time is equivalent across devices (512 ± 37 min). n = 64. *p < 0.0001. [a]Comparison with PSG. [b]Comparison with UP. SOL, sleep onset latency, TST, total sleep time, WASO, wake after sleep onset, SE, sleep efficiency.

Table 4—Results of nonparametric paired comparisons of sleep characteristics between PSG and MotionX.

	Devices	
Sleep Characteristics	**PSG**	**MotionX 24/7**
SPT (min)	519 (486–538)	535 (515–554)*
SOL (min)	26 (12–50)	14 (7–24)*
TST (min)	402 (351–460)	508 (472–535)*
WASO (min)	63 (30–111)	0 (0–14)*
SE (%)	80 (71–88)	97 (92–99)*

Data are expressed as median (interquartile range). n = 64. *p < 0.0001. SPT, sleep period time; SOL, sleep onset latency; TST, total sleep time; WASO, wake after sleep onset; SE, sleep efficiency.

Table 5—Actigraphy and UP sensitivity (ability to assess true sleep), specificity (ability to assess true wake), and accuracy (ability to assess both sleep and wake) against PSG.

	n	Sensitivity	Specificity	Accuracy
Actiwatch2	75	0.93	0.63	0.87
UP	67	0.92	0.66	0.86

the difference in SPT, MotionX 24/7 did not accurately reflect total time asleep or awake when compared to PSG. SOL and WASO were significantly underestimated (12 and 63 min, respectively), resulting in significantly longer TST and greater SE (106 min and 17%, respectively), when assessed by the MotionX 24/7.

Sensitivity, Specificity, and Accuracy

Results for the epoch-by-epoch comparisons of the UP and actigraphy against PSG are presented in **Table 5**. The overall sensitivity, specificity, and accuracy were similar across both devices.

Level of Agreement between PSG, Actiwatch2, and UP

Figure 1 shows the Bland-Altman plots for comparisons of SOL, TST, WASO, and SE between PSG and actigraphy. **Figure 2** shows the comparisons for PSG and UP.

The Bland Altman plot confirms the direction of the bias (or difference) seen in **Table 2**, with actigraphy underestimating SOL (21.4 min) and overestimating TST (17.3 min), WASO (4.1 min), and SE (3.4%). The UP device overestimated TST (9.1 min) and SE (1.8%), and underestimated WASO (9.4 min). The limits of agreement showed a smaller range for the UP device compared to actigraphy, indicating less random fluctuations around the mean. The error margin for UP, but not actigraphy, remains stable irrespective of variability in PSG outcomes. As can be seen in **Figure 1**, the longer the SOL (1A) and WASO (1C), the shorter the TST (1B), and the poorer the sleep efficiency (1D), the greater the difference between PSG and actigraphy. This trend is not seen in the PSG and UP plots.

The proportion of participants where the difference from PSG was outside clinically acceptable limits for actigraphy assessed sleep parameters were 50%, 47%, and 55% for TST (> 30 min), WASO (> 30 min), and SE (> 5%), respectively. The proportions for UP were 36% for TST, 41% for WASO, and 47% for SE.

Effect of Sleep Disordered Breathing Severity

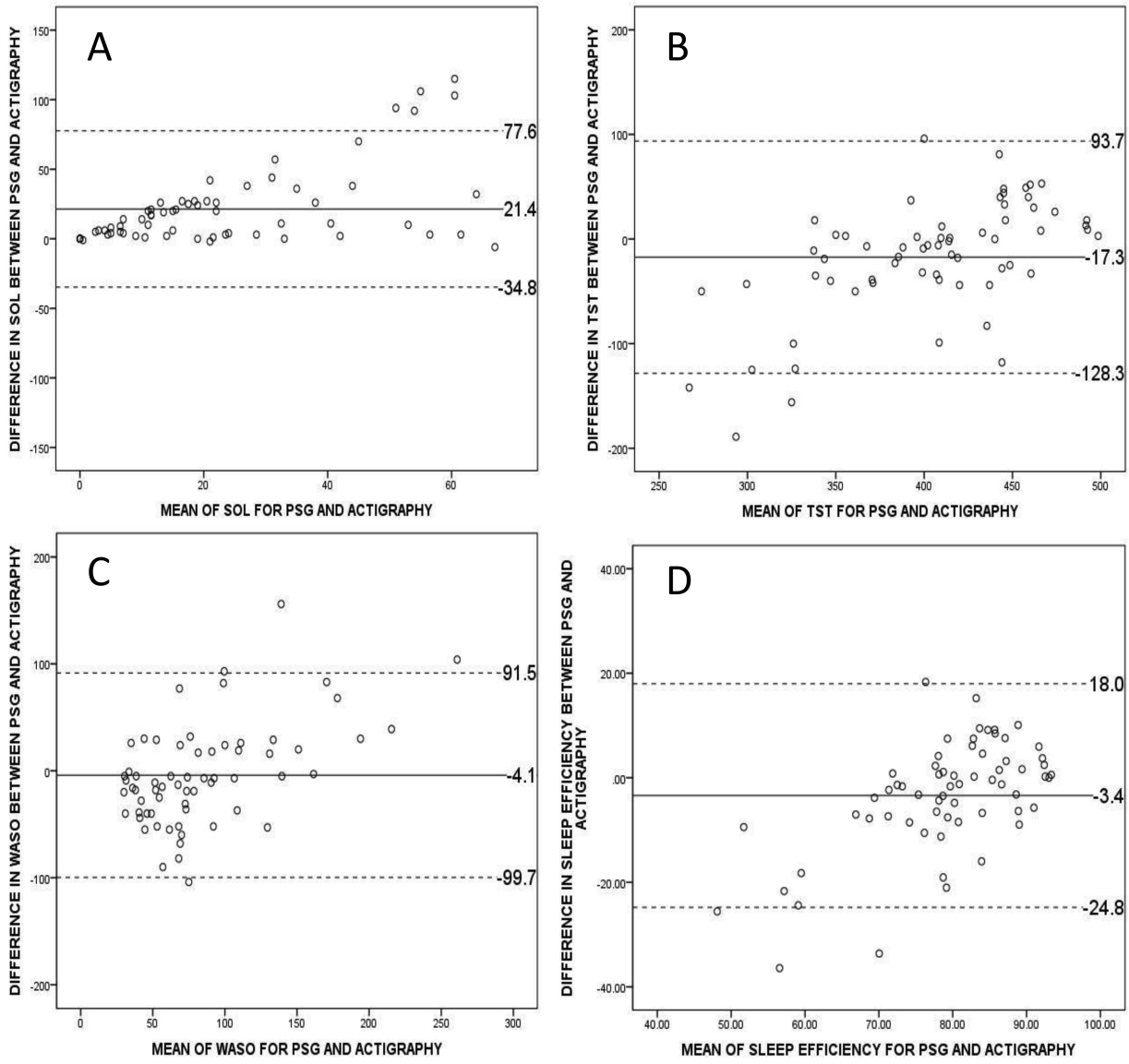
Univariate ANCOVA, controlling for age, showed no difference in the level of bias in the UP device, compared to PSG, in TST, WASO, or SE with increasing severity. There was a difference in the bias of SOL, with the UP device significantly underestimating SOL in children with primary snoring (n = 18, mean difference = 9.7 min, limits of agreement = 73.1, −53.7) compared to children with mild OSA (n = 28, mean difference = −3.9 min, limits of agreement = 11.3, −19.1) and moderate-severe OSA (n = 18, mean difference = −4.6 min, limits of agreement = 17.8, −27.0; $F_{2, 60} = 4.3$, p = 0.02).

Effect of Age

Univariate ANCOVA, controlling for ORDI, revealed a significant difference in the level of bias of the UP device, compared to PSG, in TST, WASO, and SE with increasing age. TST was significantly underestimated in preschool aged children (n = 21, mean difference = 25.3 min, limits of agreement = 80.9, −30.3) compared to primary school (n = 27, mean difference = −19.6 min, limits of agreement = 52.4, −91.6) and adolescents (n = 16, mean difference = −36.6 min, limits of agreement = 80.6, −153.8, $F_{2, 60} = 12.5$, p < 0.001), for which UP overestimated TST.

The UP device significantly overestimated WASO in preschool aged children (mean difference = −21.1, limits of agreement = 27.9, −70.1) compared to primary school (mean difference = 18.2, limits of agreement = 93.0, −56.6)

Figure 1—Bland-Altman plots for PSG versus actigraphy.



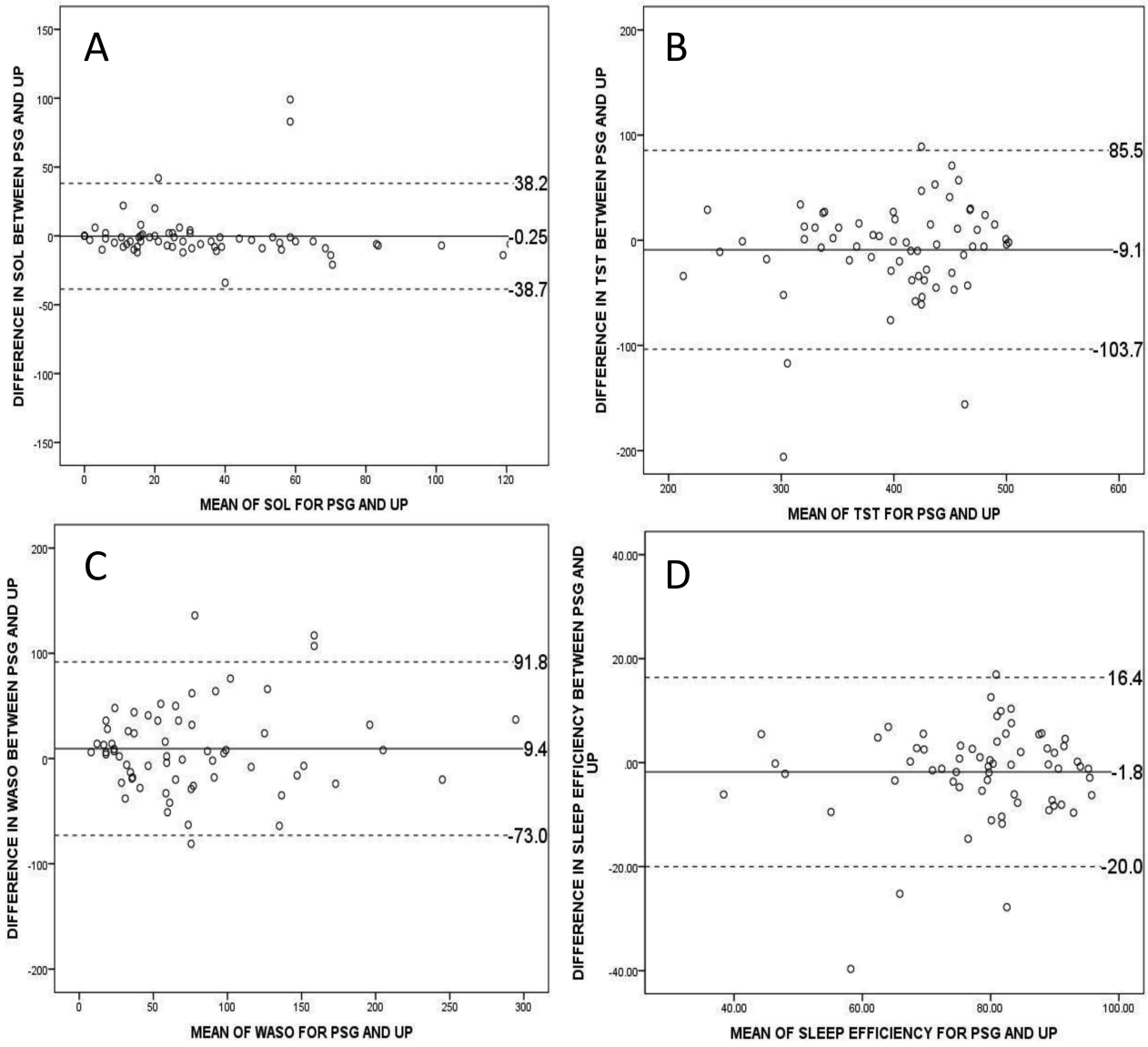
Bland-Altman plots for PSG versus actigraphy comparing sleep onset latency (SOL) **(A)**, total sleep time (TST) **(B)**, wake after sleep onset (WASO) **(C)**, and sleep efficiency **(D)**. The solid line indicates the mean of the differences or bias, and the dashed lines indicate the limits of agreement (mean difference ± 2 standard deviation) or random fluctuation around the mean.

and adolescents (mean difference = 34.3, limits of agreement = 119.3, −50.7, $F_{2,60}$ = 13.1, p < 0.001). Sleep efficiency was also significantly underestimated in preschool aged children (mean difference = 4.7%, limits of agreement = 15.1, −5.7) compared to primary school (mean difference = −3.1%, limits of agreement = 9.7, −17.1) and adolescents (mean difference = −7.4%, limits of agreement = 15.8, −30.6, $F_{2,60}$ = 12.8, p < 0.001). There were no differences between the age groups in the level of bias for SOL.

DISCUSSION

This study adds to research comparing commercially available wrist-based accelerometer devices, and to our knowledge, is the first to compare a smartphone application against PSG. The results of this study show that the wrist-based UP can assess sleep and wake to an accuracy of 86% when compared to PSG. Consistent with the few previous studies validating commercial accelerometer devices,[18–20] the UP showed greater sensitivity, accurately assessing sleep 92% of the time, than specificity, accurately assessing wake only 66% of the time. In the current study, the sensitivity and specificity results were equivalent to actigraphy, as measured with the Actiwatch2 device, when data were assessed at the medium wake threshold (93% sensitivity, 63% specificity, activity counts = 40). Interestingly, an examination of the profile of the Bland Altman scatter plots revealed the UP to be more consistent than actigraphy in assessing sleep parameters when compared to PSG. In all measures assessed, the Actiwatch2

Figure 2—Bland-Altman plots for PSG versus UP.



Bland-Altman plots for PSG versus UP comparing sleep onset latency (SOL) **(A)**, total sleep time (TST) **(B)**, wake after sleep onset (WASO) **(C)**, and sleep efficiency **(D)**. The solid line indicates the mean of the differences or bias, and the dashed lines indicate the limits of agreement (mean difference ± 2 standard deviation) or random fluctuation around the mean.

showed increasing discrepancy to PSG the shorter and more fragmented the sleep. This trend was not observed in the UP device, with equivalent margins of error occurring across all participants, regardless of their sleep duration or level of fragmentation. However, in both devices, comparisons to PSG outcomes saw differences outside clinically accepted limits in up to half of the subjects, indicating that there is unlikely to be a clinical advantage to the UP device.

In contrast to the UP wrist based device, the smartphone application (MotionX 24/7) did not accurately reflect any sleep parameter as measured by PSG. SOL and WASO were significantly shorter, TST was significantly longer, and SE significantly greater when assessed by MotionX 24/7 compared to PSG. These results may have been due to the placement of the phone, which was under the bottom bedsheet, near the right shoulder. Although this placement is recommended by the developer, this position may have resulted in the device being unable to pick up subtle movements or arousals. The developer also suggests attaching the phone to the upper arm while asleep, however this would be impractical, particularly in young children due to the size of their arms. The comparisons are also likely to be more accurate if epoch-by-epoch matching were possible. In the current study, the recording time for the MotionX 24/7 was 16 minutes longer on average than PSG. In addition, the epoch length recorded was 5 minutes compared to 1 minute, which would have substantially decreased the ability of MotionX 24/7 to report short awakenings (< 5 min). Nonetheless, the discrepancies between the two devices cannot be completely accounted for by these methodological differences.

The results of the current study add to the recent work by others comparing the Fitbit[18] and the UP[19] to PSG in children and adolescents. Zambotti et al.[19] showed similar levels of bias between the UP device and PSG in a cohort of healthy adolescents (mean age 15.8 ± 2.5y) recruited from the community. In that study, UP overestimated TST and SE by 10 minutes and 1.9%, respectively, and underestimated WASO by 9.3 minutes. SOL was also not different in that study. Meltzer and colleagues conducted a study of 63 children aged 3–17 years, and showed that the Fitbit, when set at normal mode, had an accuracy of 84% against PSG.[18] Sensitivity and specificity were both slightly lower in that study compared to the current study (sensitivity 87% versus 92%; specificity 52% versus 66%). Also significant differences were found between Fitbit and PSG for TST, WASO, and SE, which were not observed in the current comparisons between UP and PSG.[18] The differences between that study and the current one may be due to methodological differences or potential differences in scoring algorithms between the Fitbit and UP. The proprietary algorithms were not available for either of these studies.

While the large proportion of subjects with differences between UP and PSG outside clinically acceptable thresholds suggests caution if using this device as a diagnostic tool for sleep disorders relating to arousal or fragmentation, this study shows that UP is comparable to Actiwatch2 when assessing sleep and wake. Actigraphy does have advantages over the UP in that the software provides for greater flexibility in the amount of data collected and extracted for analysis. Currently, minute-by-minute sleep-wake data must be extracted manually from the UP by scrolling a finger along the hypnogram on screen, and only one day of data can be viewed at a time. Additionally, there is no avenue for altering activity thresholds or assessing other aspects of the environment, such as light. Finally, the UP was easier than the Actiwatch2 for the children to remove during sleep, as evidenced by the 13% versus 4% missing data, respectively. Regardless, this study suggests the UP may be of use in the clinical setting, particularly if one is interested in sleep-wake patterns. Further studies are needed to validate commercial devices over longer periods of time in the home setting to support this.

The secondary analyses revealed that the UP device performed consistently across all severities of SDB, with no difference in the level of bias observed between children with Primary snoring, mild OSA or moderate-severe OSA in TST, WASO, or SE. There was, however, a significant effect of age, with the direction of bias changing from preschool aged children to both primary school-aged children and adolescents. In preschool children, the UP underestimated TST and SE, and overestimated WASO. However, in both primary school-aged children and adolescents, TST and SE were overestimated and WASO underestimated, with adolescents showing greater bias (although not statistically significant) than primary school aged children. This is in contrast to Meltzer et al.,[18] who show that although there is a significant difference between PSG and Fitbit at each developmental age, the direction of this difference is consistent with WASO underestimated in the normal mode, and TST and SE overestimated. The opposite was found when PSG was compared to Fitbit in the sensitive mode. While most actigraphy validation studies discuss an overestimation of WASO in infants and children,[15,16,27] the direction of error in other pediatric[14] and adult studies is not as consistent.[28–30] It may be that as children age, the level of movement during sleep not associated with an awakening decreases. Alternatively, older children may be more likely to lie still when awake than younger children, either of which will result in the switch in the direction of bias. These developmental differences in the efficacy of accelerometer devices are yet to be elucidated, however are worthy of consideration when using these devices in clinical practice.

This study had some limitations. First, the cohort consisted solely of a clinical sample of children with suspected OSA, both with and without comorbidities. It is well known that children with OSA have increased arousal and sleep disruption compared to non-snoring children. The results may not be generalizable to a non-snoring pediatric population, where patterns of arousal and nocturnal awakenings are likely to be different. However, the similarity between the current results and those reported by Zambotti et al.[19] is encouraging. Second, collapsing the PSG data from 30-second epochs to 1-minute epochs will have overinflated the amount of wake actually scored during the sleep period, potentially overestimating specificity. Finally, data were collected for only one night in the sleep laboratory. An early study has shown that five or more nights are needed to ensure reliability of actigraphy assessments of sleep in children and adolescents,[31] and first night effects of laboratory PSG can influence sleep outcomes.[32] As such, this study cannot make inferences to a particular device's ability to assess sleep and wake over long periods of time or in more natural environments. Further research, comparing actigraphy and commercial devices over longer periods of time would be beneficial to truly understand the efficacy of these devices.

CONCLUSION

The findings of this study suggest that a wrist-based accelerometer, such as the UP by Jawbone, could be used in a similar fashion to actigraphy, such as that measured by Actiwatch2, for assessing sleep-wake patterns in children with sleep disordered breathing. When placed on the bed, accelerometer smartphone applications, such as the MotionX 24/7, do not appear to accurately reflect sleep duration or quality, and should be considered carefully before using in a clinical or research setting.

ABBREVIATIONS

ORDI, obstructive respiratory disturbance index
PSG, polysomnography
SDB, sleep disordered breathing
SE, sleep efficiency
SOL, sleep onset latency
TST, Total sleep time
WASO, wake after sleep onset

REFERENCES

1. Ringli M, Huber R. Developmental aspects of sleep slow waves: linking sleep, brain maturation and behavior. Prog Brain Res 2011;193:63–82.
2. Astill RG, Van der Heijden KB, Van Ijzendoorn MH, Van Someren EJ. Sleep, cognition, and behavioral problems in school-age children: a century of research meta-analyzed. Psychol Bull 2012;138:1109–38.
3. Biggs SN, Vlahandonis A, Anderson V, et al. Long-term changes in neurocognition and behavior following treatment of sleep disordered breathing in school-aged children. Sleep 2014;37:77–84.
4. Gregory AM, O'Connor TG. Sleep problems in childhood: a longitudinal study of developmental change and association with behavioral problems. J Am Acad Child Adolesc Psychiatry 2002;41:964–71.
5. Jackman AR, Biggs SN, Walter LM, et al. Sleep-disordered breathing in preschool children is associated with behavioral, but not cognitive, impairments. Sleep Med 2012;13:621–31.
6. Armitage R, Hoffmann RF, Emslie GJ, Weinberg WA, Mayes TL, Rush AJ. Sleep microarchitecture as a predictor of recurrence in children and adolescents with depression. Int J Neuropsychopharmacol 2002;5:217–28.
7. Gregory AM, Van der Ende J, Willis TA, Verhulst FC. Parent-reported sleep problems during development and self-reported anxiety/depression, attention problems, and aggressive behavior later in life. Arch Pediatr Adolesc Med 2008;162:330–5.
8. Anstead M. Pediatric sleep disorders: new developments and evolving understanding. Curr Opin Pulm Med 2000;6:501–6.
9. Marcus CL. Sleep-disordered breathing in children. Am J Respir Crit Care Med 2001;164:16–30.
10. Marcus CL, Traylor J, Biggs SN, et al. Feasibility of comprehensive, unattended ambulatory polysomnography in school-aged children. J Clin Sleep Med 2014;10:913–8.
11. Meltzer LJ, Montgomery-Downs HE, Insana SP, Walsh CM. Use of actigraphy for assessment in pediatric sleep research. Sleep Med Rev 2012;16:463–75.
12. Sadeh A. The role and validity of actigraphy in sleep medicine: an update. Sleep Med Rev 2011;15:259–67.
13. Lovato N, Gradisar M, Short M, Dohnt H, Micic G. Delayed sleep phase disorder in an Australian school-based sample of adolescents. J Clin Sleep Med 2013;9:939–44.
14. Hyde M, O'Driscoll DM, Binette S, et al. Validation of actigraphy for determining sleep and wake in children with sleep disordered breathing. J Sleep Res 2007;16:213–6.
15. Meltzer LJ, Walsh CM, Traylor J, Westin AM. Direct comparison of two new actigraphs and polysomnography in children and adolescents. Sleep 2012;35:159–66.
16. Meltzer LJ, Westin AM. A comparison of actigraphy scoring rules used in pediatric research. Sleep Med 2011;12:793–6.
17. Sadeh A, Sharkey KM, Carskadon MA. Activity-based sleep-wake identification: an empirical test of methodological issues. Sleep 1994;17:201–7.
18. Meltzer LJ, Hiruma LS, Avis KT, Montgomery-Downs HE, Valentin J. Comparison of a commercial accelerometer with polysomnography and actigraphy in children and adolescents. Sleep 2015;38:1323–30.
19. de Zambotti M, Baker FC, Colrain I. Validation of sleep-tracking technology compared with polysomnography in adolescents. Sleep 2015;38:1461–8.
20. Montgomery-Downs HE, Insana SP, Bond JA. Movement toward a novel activity monitoring device. Sleep Breath 2012;16:913–7.
21. Berry RB, Budhiraja R, Gottlieb DJ, et al. Rules for scoring respiratory events in sleep: update of the 2007 AASM Manual for the Scoring of Sleep and Associated Events. Deliberations of the Sleep Apnea Definitions Task Force of the American Academy of Sleep Medicine. J Clin Sleep Med 2012;8:597–619.
22. Nixon GM, Hyde M, Biggs SN, Walter LM, Horne RS, Davey MJ. The impact of recent changes to the respiratory scoring rules in pediatrics. J Clin Sleep Med 2014;10:1217–21.
23. 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.
24. Tabachnick BG, Fidell LS. Using multivariate statistics, 4th ed. Boston, MA: Allyn and Bacon, 2001.
25. Bland JM, Altman DG. Comparing methods of measurement: why plotting difference against standard method is misleading. Lancet 1995;346:1085–7.
26. Werner H, Molinari L, Guyer C, Jenni OG. Agreement rates between actigraphy, diary, and questionnaire for children's sleep patterns. Arch Pediatr Adolesc Med 2008;162:350–8.
27. Insana SP, Gozal D, Montgomery-Downs HE. Invalidity of one actigraphy brand for identifying sleep and wake among infants. Sleep Med 2010;11:191–6.
28. Hauri PJ, Wisbey J. Wrist actigraphy in insomnia. Sleep 1992;15:293–301.
29. 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.
30. Pollak CP, Tryon WW, Nagaraja H, Dzwonczyk R. How accurately does wrist actigraphy identify the states of sleep and wakefulness? Sleep 2001;24:957–65.
31. Acebo C, Sadeh A, Seifer R, et al. Estimating sleep patterns with activity monitoring in children and adolescents: how many nights are necessary for reliable measures? Sleep 1999;22:95–103.
32. Agnew HW Jr., Webb WB, Williams RL. The first night effect: an EEG study of sleep. Psychophysiology 1966;2:263–6.

ACKNOWLEDGMENTS

Contributors Statement: Elicia Toon collected all data, conducted the data analysis, drafted the initial manuscript, and approved the final manuscript as submitted. Samantha L. Hollis participated in data entry and analysis, reviewed and revised the manuscript, and approved the final manuscript as submitted. Margot J. Davey, Gillian M. Nixon, and Rosemary S.C. Horne participated in the conceptualization and design of the study, provided funding and resources for the study, critically review the manuscript, and approved the final manuscript as submitted. Sarah N. Biggs participated in the conceptualization and design of the study, provided supervision and project management of data collection, conducted the final data analysis, critically reviewed and redrafted the manuscript, and approved the final manuscript as submitted.

SUBMISSION & CORRESPONDENCE INFORMATION

Submitted for publication May, 2015
Submitted in final revised form September, 2015
Accepted for publication September, 2015
Address correspondence to: Dr. Sarah Biggs, The Ritchie Centre, Hudson Institute of Medical Research, Level 5 Monash Medical Centre, 246 Clayton Road, Clayton, Vic 3168, Australia; Tel: +61 3 85722825; Fax: +61 3 95946811; Email: sarah.biggs@monash.edu

DISCLOSURE STATEMENT

This was not an industry supported study. The Hudson Institute is supported by the Victorian Government's Operational Infrastructure Support Program. The authors have indicated no financial conflicts of interest.