ELSEVIER

Sleep Medicine 5 (2004) 541–550

SLEEP MEDICINE

www.elsevier.com/locate/sleep

Original article

Wrist actigraphy in estimation of sleep and wake in intellectually disabled subjects with motor handicaps

Maija-Liisa Laakso[a,b,*], Lea Leinonen[a,b], Niki Lindblom[a], Sirkka-Liisa Joutsiniemi[b], Markus Kaski[a]

[a]*Rinnekoti Foundation, Sleep Research Centre, Kumputie 1, FIN-02980 Espoo, Finland*
[b]*Neural Networks Research Centre, University of Technology, Espoo, Finland*

Received 8 March 2004; received in revised form 18 May 2004; accepted 20 May 2004

Abstract

Background and purpose: We studied the applicability of wrist actigraphy to sleep–wake estimation in patients with motor handicaps.

Patients and Methods: Concomitant polysomnographic and actigraphic recordings (16–24 h) were compared in three groups: normally moving subjects with normal sleep ($n=10$), sleep-disordered subjects without motor handicaps ($n=13$) and sleep-disordered patients with different motor disabilities ($n=16$). The motor abilities of the subjects were determined by clinical evaluation using a grading scale from 0 to 10. Their actual daily activity was calculated from the recordings as average activity scores.

Results: In the healthy subjects, the mean difference between actigraphic and polysomnographic total sleep estimation was negligible (−1 min), while in both sleep-disordered groups, sleep was highly overestimated by actigraphy. There was a significant correlation between the motor ability score and the discrepancy between actigraphy and polysomnography, but individual data points were highly scattered. A more consistent correlation was found between the average activity score/min in actigrams and the discrepancy of actigraphic with polysomnographic total sleep estimation (Spearman's $r=-0.58$, $P=0.0001$, $n=39$). When the recordings with very low average activity score were rejected from the analyses (two patients without and six with motor handicaps), the overestimation of sleep by actigraphy was reduced but it still remained in both sleep-disordered groups. The mean differences of total sleep between actigraphy and polysomnography were 72 and 121 min and the rank order correlation coefficients 0.80 and 0.71 in patients without and with motor handicaps, respectively. The median discrepancy in total sleep estimation was 6% in both sleep-disordered groups.

Conclusions: In subjects with rudimentary motor abilities, a standard actigraphy can produce a signal, which is related to the amount of sleep scored in polysomnograms. The sleep parameters obtained by the two methods are not equal, however. The inspection of actigrams is more reliable than the clinical scaling of motor abilities in predicting the applicability of wrist actigraphy.

Keywords: Actigraphy; Polysomnography; Mental retardation; Motor disorder; Sleep disorder; Validation

1. Introduction

Measurement of wrist activity is a widely used method in the estimation of sleep and wake in long-term recordings. When applied to healthy people with normal sleep, the method usually gives results which are very similar to polysomnographic (PSG) measures of sleep and wake, but the agreement is not as good if the method is used in investigating sleep-disordered subjects [1–3]. Sleep may be highly overestimated, especially in patients who lie in bed waking but motionless for long periods. This problem has been encountered not only in persons suffering from depression [4] and other insomniacs, [5] but also in hypersomnic patients [6].

Actigraphy (AGR) has also been applied to persons whose motor abilities deviate from the norm. In such cases, however, conclusions concerning sleep and wake have not usually been drawn. For instance, in patients with developmental or degenerative brain disorders the method has been

* Corresponding author. Address: Rinnekoti Foundation, Sleep Research Centre, Kumputie 1, FIN-02980 Espoo, Finland. Tel.: +358-9-8551424; fax: +358-9-8551344.

E-mail address: maija-liisa.laakso@rinnekoti.fi (M.-L. Laakso).

1389-9457/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.sleep.2004.05.002

used to get information about the rest-activity rhythm [7–11]. We have come across only one previous attempt to test the agreement between AGR and PSG in assessing sleep and wake in subjects with motor disturbances [12]. In 10 wheelchair-bound demented patients, the rank order correlation between total sleep times determined by electroencephalography and AGR was up to 0.91, suggesting that the results obtained by the two methods were related. Differences between the results obtained by the two methods were large in some individuals, and the correlation was weaker for the other tested parameters. In our clinical recordings of some intellectually disabled patients with severe motor handicaps, there has been a surprisingly good agreement between actigrams and the sleep logs kept by the nurses, although sometimes only the 'caregivers' activity' has been recorded (unpublished observations). To date, there is no knowledge of the minimum motor abilities required for gaining acceptable AGR recordings for estimating sleep and wake in handicapped patients.

In the present study, we planned to characterize the motor functions required for reliable AGR sleep estimation. When overnight PSG recordings were performed for diagnostic purposes in patients with developmental brain disorders, we graded the motor abilities of the patients and recorded wrist activity concomitantly with PSG. The rationale was to find a practical 'lowest motor ability score' that would warrant the use of AGR (e.g. in follow-up studies during possible interventions). Because this approach yielded inadequate results, we searched for additional predicting features from the AGR recording itself.

2. Subjects and methods

2.1. Subjects

The data consisted of overnight PSG and AGR recordings (16–24 h) of nine healthy volunteers with normal sleep, five intellectually normal subjects with sleep disorders (staff of the institutes), and 25 intellectually disabled subjects who were studied for suspected sleep disturbances. The subjects were divided into three groups according to their motor abilities and complaints of sleep problems (Table 1). Group I consisted of normally moving subjects who reported no sleep problems. One patient was recorded due to challenging daytime behaviour and no sleep disorder was found; he was included in Group I. Group II consisted of subjects with sleep disorders but without motor handicaps (both staff members and patients) and Group III of patients with sleep disorders and motor handicaps. The degree of motor disability in Group III was assessed by clinical observations and neurological examination, as shown in Table 2. Well-sleeping subjects with motor handicaps were not found among the recorded patients.

The causes of the patients' brain disorders were diverse: pre or postnatal brain damage in six (infections, contusions, asphyxia); known genetic disorders in nine (aspartylglucosaminuria in two patients; partial trisomy 12, Angelman, Pallister-Killian, Rubinstein-Taybi, Smith-Magenis, Sotos and Wolf-Hirschhorn syndromes in each of the others); unknown familial disorders in three and an unknown etiology in seven subjects.

Table 1
Characteristics of the study groups

	Group I	Group II	Group III
Sleep	Normal	Disordered	Disordered
Motor abilities	Normal	Normal	Limited
N	10	13	16
Females/males	3/7	8/5	5/11
Age (y, mean ± SD)	28 ± 10	38 ± 14	36 ± 13
Intellectually normal/disabled	9/1	5/8	0/16
Blind	0	1 (8%)	4 (25%)
Epilepsy	0	3 (23%)	12 (75%)
Motor ability score[a] (mean ± SD)			
Locomotion	5 ± 0	5 ± 0	1.9 ± 1.7
Arm movements	5 ± 0	5 ± 0	2.4 ± 1.4
Total	10 ± 0	10 ± 0	4.3 ± 2.3
Sleep disorder complaints[b]			
Frequent daytime sleep	–	5 (38%)	8 (50%)
Frequent short awakenings	–	6 (46%)	6 (38%)
Long wake periods at night	–	7 (54%)	10 (62%)

[a] For definition see Table 2.
[b] Questionnaire-based main symptoms: frequent daytime sleep = on 5 days/week or more; frequent short awakenings = on scale of minutes, almost every night ≥ 3 awakenings/night; long wake periods = on scale of hours, at least on 2–3 nights/week. Some subjects had several kinds of complaints.

Table 2
Definition of motor ability scores

	Score
Locomotion	
Able to walk independently	5
Unable to walk but able to move independently in some other way, e.g. on buttocks or by crawling, able to leave the bed independently	4
Unable to change place independently but able to walk with help	3
Unable to change place even with help but able to get up to a sitting or crawling position	2
Unable to get up but able to change the lying position independently	1
Unable to change position in bed	0
Arm movements	
Normal (or almost normal) voluntary arm movements	5
Limited voluntary arm movements but manages some everyday tasks, e.g. able to eat or undress independently	4
Unable to eat independently but able to display some goal-directed arm movements, e.g. for toys or caretakers	3
No voluntary arm movements but displays stereotypic movements (e.g. tapping, swaying, athetotic movements) while waking	2
No voluntary or stereotypic arm movements but able to hold things or let them fall	1
No arm movements	0

The sum of 'locomotion' and 'arm movement' score was used in the study.

The study plan was accepted by the local ethical committee. Written informed consent was obtained from the intellectually normal subjects and from the official representatives of the intellectually disabled subjects.

2.2. Polysomnography and visual scoring

The long-term ambulatory polygraphs were performed with Embla (Flaga, Iceland) at Rinnekoti Centre for Intellectually Disabled. The healthy volunteers were recorded at home or in the sleep laboratory, according to their wish, and the patients in the residential quarters of the Centre. The volunteers themselves kept logs of their activities while waking, the patients were continuously observed and the observer kept detailed logs. For visual sleep scoring, two electro-oculography (EOG) leads, C3/A2 electroencephalography (EEG) lead and submental electromyography (EMG) were recorded. The recordings also included frontal, parietal and temporal EEG. Respiratory and movement disorders were monitored with additional sensors. EOG and EEG were sampled at 100 Hz, EMG and ECG at 200 Hz using 16-bit resolution.

The 30 s scoring epochs were labelled as wake (W), sleep stages 1–4 (S1, S2, S3, S4), rapid eye movement sleep (REM) or movement time (MT). The recordings were scored according to the standard guidelines [13] with the following exceptions. Each epoch was scored irrespective of the neighbouring epochs, and MT (at least 15 s segments characterized by movement artefacts and an increase in EMG) was also noted across epochs (the epoch with the longer movement duration was labelled). Some subjects with brain pathologies had no alpha during waking or vertex waves or sigma during sleep. Nevertheless, the general course of EEG events was similar in subjects with normal and pathological EEG. After the cessation of blinking and movements, the beginning of S1 was detected by the abrupt change of relaxed waking EEG and the appearance of low amplitude theta. The increase in the proportion of 0.5–3 cps activity exceeding a peak-to-peak threshold value was used for the differentiation among S2, S3 (more than 20% of time) and S4 (more than 50% of time). For some patients the threshold had to be set lower than 75 μV, recommended for healthy subjects [13]. After an arousal or awakening there was a period of REM characterized by S1-type EEG, low or phasic EMG and bursts of rapid eye movements. In some patients there was no discernible EMG activity during REM. REM was interrupted by wakefulness more frequently in patients than healthy subjects. Some patients with sleep disorders had short (<15 s) arousals which were not taken into account in the present scoring; recurrent arousals usually resulted in long periods of sleep scored alternatingly as S1 or S2.

2.3. Actigraphy and data analysis

Activity was measured with a piezo-electric accelerometer (Actiwatch, Cambridge Neurotechnology Ltd, Cambridge, UK), which records all movement exceeding 0.05 g in all directions. The activity counts stored in the memory unit of the device are produced by integration of intensity, amount and duration of movement. The signal was sampled at 32 Hz using 3–11 Hz bandpass filtering. The subjects who moved normally wore the device on their non-dominant wrist and the subjects with motor handicaps on the more mobile wrist if there was a difference between the arms. Six watches were randomly used in each group. Data were saved in 1 min epochs. The Actiwatch algorithm calculates a final activity score for each epoch by considering the activity counts in the neighbouring epochs as follows: $A = 0.04E_{-2} + 0.20E_{-1} + E_0 + 0.20E_{+1} + 0.04E_{+2}$, where A is the final activity score for epoch 0, E_0, the original activity counts for epoch 0, and the other epochs are the activity counts within 1 and 2 min before and after epoch 0.

The night time AGR sleep parameters were determined automatically by Actiwatch Sleep Analysis software (version 4.15, Cambridge Neurotechnology Ltd) after giving the bed and get up times to the program. The software allows sleep–wake scoring at three sensitivity levels. Medium sensitivity (score ≥ 40/min = 'awake') was used if the average daily activity score was >100/min, and high sensitivity setting (score ≥ 20/min = awake) if it was <100/min (Section 4). For automatic determination of sleep onset following bedtime, the algorithm selects the first consecutively recorded data of at least 10 min of immobility, with no more than 1 epoch of movement within that time. For determination of sleep end, the algorithm selects a corresponding 10 min sequence of immobility preceding get up time.

To determine the amount of daytime sleep, Actiwatch Nap Analysis algorithm was applied to the recording periods before bed time and after get up time. If the average daily activity score was >100/min, the sensitivity was set to activity score ≤ 10/min defined as 'sleep'. If the average activity was <100/min, the sensitivity was set to score 0. The minimum duration of daytime sleep sequence taken into account was 10 min.

2.4. Definitions

The sleep parameters determined for comparisons between AGR and PSG were: *total sleep time* (PSG: sum of all sleep epochs during the recording period; AGR: sum of actual sleep in Sleep Analysis and naps in Nap Analysis), *night sleep* (sum of all sleep epochs during the period in bed), *daytime sleep* (sum of all sleep epochs during the period outside of bed), *latency* (sleep onset latency after bedtime), and *efficiency* (percentage of night sleep during the period in bed).

Motor ability score was used to define the clinically evaluated motor disturbances of the patients (Table 2). *Average activity score/min* was calculated from the AGR recordings and described the actual activity of the subjects.

2.5. Calculations and statistics

Non-parametric tests were applied in comparisons of sleep variables among the groups (non-normal distributions and/or non-homogeneity of variances in many instances): Kruskal–Wallis test followed by Dunn's test in comparisons of several groups, and Wilcoxon's paired or Mann–Whitney's unpaired test in comparisons of two groups.

The agreement between AGR and PSG sleep parameters was evaluated by calculating the means and standard deviations of the differences between the parameters obtained by the two methods [14]. Spearman's rank order correlation coefficients were also calculated to evaluate whether the parameters obtained by the two methods were related.

Discrepancy between PSG and AGR assessment of total sleep was calculated in percentage as follows: 100× (sleepAGR − sleepPSG)/recording period, where sleepAGR is the total sleep assessed automatically in AGR, and sleepPSG is the total sleep scored visually from PSG recordings. In addition, corresponding percentage discrepancies were calculated separately for the periods in bed and out of bed. Spearman's rank order correlation coefficients were used for describing relationships between discrepancy% and motor ability score, discrepancy% and average activity score/min in AGR, and between discrepancy% and selected PSG sleep parameters. Partial correlation coefficients were also calculated to evaluate the share of actual activity in the correlations [15].

3. Results

3.1. Polysomnographic sleep parameters

In the group of healthy volunteers without sleep disorder complaints (Group I), the PSG parameters did not deviate much from the values presented for normal sleep in handbooks, while in the other two groups with anamnestic sleep disturbances, there was a great dispersion of the parameters (Table 3). The pathologies tended to be qualitatively similar in the two groups with sleep disorders, but the deviations of the parameters were more pronounced in the patients with motor handicaps.

Because both long- and short-sleep patterns were found in the two sleep-disordered groups, the medians of total sleep time or night sleep period did not differ significantly among the groups (Table 3). As expected, the subjects with motor handicaps spent more time in bed than the normally moving subjects. This may explain the finding that sleep onset latency was very long in many subjects of Group III. The median number of night time awakenings did not differ significantly among the groups, but long-lasting wake periods were more frequent in Groups II and III than in

Table 3
Polysomnographic parameters (medians and ranges) in the study groups

	Group I	Group II	Group III	Kruskal–Wallis
Sleep Motor abilities	Normal Normal	Disordered Normal	Disordered Limited	Test *P*
Total sleep (h)	8.4 (7.0–9.7)	6.5 (4.0–14.2)	7.5 (0.5–13.2)	NS
Night sleep				
In bed (h)	9.5 (7.7–11.2)	9.6 (6.6–11.2)	11.8 (8.9–14.2)**$^{\#\#\#}$	<0.0001
Latency (min)	14 (1–73)	48 (8–258)	121 (7–606)**	0.0033
Sleep period (h)	9.0 (7.5–10.0)	7.8 (4.1–11.1)	8.6 (1.0–13.3)	NS
N of awakenings				
0.5–4.5 min	19 (3–36)	26 (2–55)	13 (3–93)	NS
5 min or longer	1 (0–4)	2 (0–12)	5 (0–10)**	0.0058
WASO (min)	31 (4–70)	54 (2–318)	120 (6–369)**	0.0082
Actual sleep (h)	8.3 (7.0–9.7)	6.2 (3.1–10.0)*	5.9 (0.5–9.0)*	0.0120
Efficiency (%)	87 (77–98)	77 (28–94)	51 (5–80)***	0.0002
Daytime sleep				
N of naps	0 (0–1)	1 (0–3)	1 (0–6)	[0.0599]
Sleep epochs (min)	0 (0–26)	3 (0–250)	31 (0–248)	[0.0566]
Percent of total sleep	0 (0–5)	1 (0–46)	9 (0–31)	[0.0539]
Sleep stages				
S1%	12 (8–25)	35 (12–75)***	23 (3–63)*	0.0004
S2%	43 (33–49)	31 (10–49)	28 (20–55)**	0.0092
S3 + S4%	19 (12–34)	15 (0–59)	30 (5–44)$^{\#\#}$	0.0057
REM%	20 (12–31)	12 (4–20)**	13 (3–27)*	0.0046
MT epochs N	23 (13–44)	12 (4–45)	8 (3–46)*	0.0189

*Different from Group I, $P<0.05$; **$P<0.01$; ***$P<0.001$; $^{\#\#}$different from Group II, $P<0.01$; $^{\#\#\#}P<0.001$ (Dunn's test). Total sleep, sum of night and daytime sleep epochs; sleep period, the period from the first S1 epoch after bedtime to the first W epoch not followed by any sleep epochs preceding get up time; WASO, wake after sleep onset; actual sleep, sum of sleep epochs during the period in bed.

Table 4
Differences and correlations between the sleep parameters assessed by actigraphy (AGR) and polysomnography (PSG) in the study groups

	Group I	Group II	Group III
Sleep	Normal	Disordered	Disordered
Motor abilities	Normal	Normal	Limited
N	10	13	16
Differences with standard deviations (AGR–PSG), times in minutes			
Total sleep	−1±41	+93±82	+248±290
Night sleep	−17±28	+60±68	+184±158
Daytime sleep	+16±28	+33±58	+65±145
Latency	−6±7	−48±62	−152±194
Efficiency%	−3±5	+10±11	+25±22
Spearman's rank order correlation coefficients (P-values)			
Total sleep	0.77 (0.0046)	0.87 (<0.0001)	0.26 (NS)
Night sleep	0.88 (0.0008)	0.82 (0.0003)	0.29 (NS)
Daytime sleep[a]	–	0.88 (<0.0001)	0.62 (0.0055)
Latency	0.82 (0.0029)	0.73 (0.0022)	0.17 (NS)
Efficiency%	0.77 (0.0063)	0.91 (<0.0001)	0.19 (NS)

[a] Only one subject in Group I had daytime sleep in polysomnography.

Group I, resulting in longer wake time after sleep onset (WASO), especially in Group III. Actual night sleep was significantly shorter in Groups II and III than in Group I. Due to the long time in bed, long sleep onset latency and long WASO, the sleep efficiency was low in most subjects of Group III, and the efficiency in Group II tended to be lower than in Group I. Although many sleep-disordered subjects slept during the day, and some of them for several hours, the differences of the daytime sleep parameters were not quite significant among the groups. The shares of the scored sleep stages varied greatly according to the types of sleep disturbances. In both sleep-disordered groups there were several subjects with a great share of stage 1 and a small share of REM sleep.

3.2. Actigraphic sleep parameters compared with polysomnography

In the group of normally moving subjects with normal sleep, automatic analysis of AGR suggested somewhat less night sleep and more daytime sleep than scored in the PSG recordings (Table 4). Thus, the mean difference in total sleep time assessed with the two methods was negligible. In both sleep-disordered groups, the AGR analysis overestimated both night and daytime sleep. Most of the overestimation of night sleep was caused by underestimation of sleep onset latency. The mean differences of total sleep determined by AGR and PSG were 1.5 and 4 h in Groups II and III, respectively.

The correlation between AGR and PSG parameters was significant in Groups I and II, while in Group III a significant correlation was found only in the duration of daytime sleep (Table 4). Thus, although the agreement of the two methods in estimating the amount of sleep in Group II was far from complete, the methods at least yielded related results. In Group III, consisting of patients with motor disabilities, the disagreement was still greater and the estimates were not even related.

In each group, there were individual recordings with poor and good agreement between the methods. The investigation was continued by searching reasons for the inter-individual differences.

3.3. Percentage discrepancy between actigraphy and polysomnography

Because there was some variation in the duration of recordings and considerable inter-individual variation in the periods in bed, we calculated relative discrepancy values for comparisons. Table 5 shows that the overestimation of total sleep by AGR was significant in the sleep-disordered groups, even when related to the total recording period, and most of the discrepancy occurred during the periods in bed. The discrepancy was extremely high in some individuals of Group III, but there were also subjects whose AGR and PSG recordings yielded similar results.

3.4. The relationship between discrepancy% and motor ability score

The original purpose of the study was to determine whether clinical evaluation of a subject's motor abilities would predict the usefulness of AGR in sleep estimation. Fig. 1A shows that when estimating total sleep there was some correlation between the discrepancy% and the motor

Table 5
Discrepancy% between polysomnographic scoring (PSG) and actigraphic assessment (AGR) of total sleep, night sleep and daytime sleep

	Group I	Group II	Group III	Kruskal–Wallis
Sleep	Normal	Disordered	Disordered	Test *p*
Motor abilities	Normal	Normal	Limited	
N	10	13	16	
Total	0 (−6 to +5)	+8 (−1 to +21)*	+13 (−6 to +72)**	0.0017
In bed	−2 (−15 to +2)	+9 (−7 to +32)*	+21 (−3 to +89)***	<0.0001
Out of bed	0 (0 to +8)	+3 (−2 to +25)	+2 (−10 to +58)	NS

Discrepancy was calculated as the difference of sleep time in AGR–sleep time in PSG, and it is given as percentage of the total recording period, period in bed and period out of bed, respectively (medians and ranges). Positive percentage means more sleep in AGR and negative percentage more sleep in PSG. Different from Group I, $P<0.05$; $^{**}P<0.01$; $^{***}P<0.001$, Dunn's test.

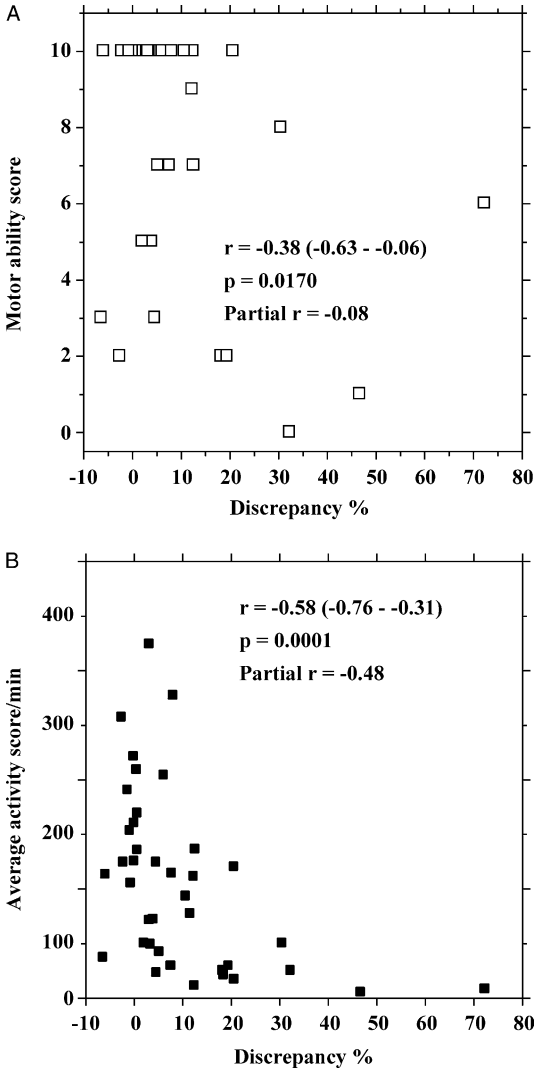

Fig. 1. Correlations between motor ability score and discrepancy of actigraphic and polysomnographic total sleep estimation (A), and between average actual activity score/min and discrepancy (B). Spearman's rank order correlation coefficients with 95% confidence intervals are given in the figure. Partial correlation coefficients were calculated by holding the third factor constant (activity score in (A), motor ability score in (B)). Number of subjects = 39. Discrepancy% was calculated as in Table 5.

ability score, but the individual data points were highly scattered. In an individual with a motor ability score 10 (normal motor abilities) the discrepancy could be as high as 20%, while in some individuals with score 2 or 3 (severely handicapped), the discrepancy remained as low as ±5%. Similar relationships were seen when the arm movement or locomotion scores (Table 2) were plotted separately against the discrepancy% (data not shown).

3.5. The relationship between discrepancy% and average activity score

Another way to evaluate the usefulness of AGR in subjects with motor handicaps is to find a predicting feature in the recording itself. With this in mind we calculated the average activity score/min in the recordings. As expected, the scores were lower in Group III than in the groups with normal mobility, especially during daytime (Table 6). However, some subjects in Group II had a very low activity score irrespective of their normal motor ability, and there were subjects in Group III with a high activity score irrespective of their limited motor abilities.

The correlation between the discrepancy% and the actual activity score/min during the whole recording period was highly significant (Fig. 1B). Similarly, a significant correlation was found between the discrepancy% and the average activity score during the period in bed (Spearman's $r = -0.36$, 95% confidence interval -0.04 to -0.61, $P = 0.0267$, $n = 39$) and during the period out of bed ($r = -0.54$, 95% confidence interval -0.26 to -0.73, $P = 0.0004$, $n = 39$). The ratio of the activity score out of bed/in bed did not correlate significantly with the discrepancy%.

To find out whether one of the two variables—the motor ability score or the actual activity score—would be a better predictor of related results between AGR and PSG, partial correlation coefficients were calculated by holding the third factor constant. When the actual activity score was held constant, the correlation between the discrepancy% and motor ability score disappeared (Fig. 1A). When the motor ability score was held constant, the correlation between the discrepancy% and actual activity score remained (Fig. 1B).

3.6. The relationships between discrepancy% and PSG sleep parameters

Because the discrepancy between AGR and PSG was greater in both sleep-disordered groups than in the group with normal sleep, we searched for PSG features that might

Table 6
Average activity score/min (medians and ranges) of actigraphic recordings in the study groups

	Group I	Group II	Group III	Kruskal–Wallis
Sleep	Normal	Disordered	Disordered	Test *P*
Motor abilities	Normal	Normal	Limited	
N	10	13	16	
Total period	144 (70–220)	114 (11–324)	33 (5–257)**	0.0021
In bed	22 (4–46)	32 (2–68)	13 (3–129)	NS
Out of bed	216 (118–362)	207 (24–491)	51 (6–359)**#	0.0015

**Different from Group I, $P < 0.01$; different from Group II; $P < 0.05$, Dunn's test.

Table 7
Spearman's rank order correlation coefficients between polysomnographic sleep parameters and discrepancy% of actigraphy with polysomnography

	Spearman's *r* (95% confidence interval)	*P*	Partial correlation coefficient
Total sleep	−0.41 (−0.65 to −0.10)	0.0094	−0.56
Night sleep			
Time in bed	0.37 (0.06 to 0.62)	0.0193	0.04
Latency	0.54 (0.26 to 0.73)	0.0004	0.47
Period	−0.32 (−0.58 to 0.01)	0.0499	−0.52
N of awakenings	−0.14 (−0.44 to 0.20)	NS	−0.13
WASO	0.17 (−0.16 to 0.47)	NS	0.02
Actual sleep	−0.44 (−0.67 to 0.14)	0.0047	−0.52
Efficiency	−0.61 (−0.78 to −0.35)	<0.0001	−0.56
Daytime sleep			
N of naps	0.04 (−0.28 to 0.36)	NS	−0.03
Sleep epochs	0.12 (−0.22 to 0.42)	NS	−0.007

Partial correlation coefficients were calculated holding constant the actual activity score/min. WASO, wake after sleep onset.

predict a high discrepancy. A long time in bed was associated with a high discrepancy% (Table 7). This correlation disappeared, however, when holding the actual activity score constant. Several coefficients suggested that the discrepancy was higher when less sleep was scored in PSG (total sleep, night sleep period and actual night sleep). A very significant correlation was found between the sleep onset latency and the discrepancy% (long latency–high discrepancy). The most significant correlation was between sleep efficiency and discrepancy% (low efficiency–high discrepancy). All these correlation coefficients remained relatively high, even when the activity score was held constant. The number of awakenings or WASO did not correlate significantly with the discrepancy%, nor did the number or duration of naps.

3.7. Minimum average activity score yielding related total sleep estimates by actigraphy and polysomnography

The above calculations suggested that the activity score might be more usable than the clinical evaluation of motor handicap in predicting the applicability of AGR in sleep estimation. In addition, sleep disturbances, especially sleep onset insomnia, seemed to be associated with a poor agreement of AGR with PSG, irrespective of the subjects' motor abilities.

Fig. 1B shows that the highest discrepancy between PSG and AGR was found in the patients with an activity score lower than 20–30/min. We therefore rejected the data from the patients with an activity score of 25 or less (two patients from Group II and six patients from Group III). Recalculations of the reduced data showed that the average actual activity in Group IIIR remained significantly lower than in the normally moving healthy control subjects, and tended to be lower than in the normally moving patients of Group IIR (Table 8). Irrespective of the different activity scores, the median discrepancy% did not differ between Groups IIR and IIIR, but was significantly higher in these groups (with sleep disturbances) compared to the well-sleeping subjects of Group I.

The range of average activity score/min was narrowed in the reduced data (Table 8) and the correlation between actual activity and discrepancy% was no more significant (data not shown). Thus, the agreement between AGR and PSG group data seemed to be substantially improved by restricting the analyses to the AGRs with an average daily activity score of more than 25/min.

The risk of discrepant estimation of total sleep by AGR and PSG in individual cases was evaluated by recalculating the mean differences and their standard deviations in the reduced groups (Fig. 2, compare with Table 4). The plot of Group I values shows that, although the mean difference was close to zero, the discrepancy was as large as an hour in some individual cases. If we expect most of the cases in population surveys to lie between the mean difference −2SD and +2SD, the expected range of differences would be −83 to +81 min in subjects with normal sleep and without motor handicaps.

In the sleep-disordered groups the overestimation of sleep by AGR persisted after rejecting the recordings with very low activity scores. The mean difference in Group II was more than an hour, and in Group III about 2 h (Fig. 2). The range between −2SD and +2SD was several hours in both groups. Thus, the sleep estimation by AGR and PSG yielded clearly different results in the subjects with sleep disorders.

Rejecting the recordings with very low activity counts from Group III increased Spearman's correlation coefficient between the AGR and PSG total sleep estimation (Fig. 2).

Table 8
Average activity score/min and discrepancy% between polysomnographic and actigraphic assessment of total sleep time in reduced study groups (IIR, IIIR)

	Group I	Group IIR	Group IIIR	Kruskal–Wallis
Sleep	Normal	Disordered	Disordered	Test *P*
Motor abilities	Normal	Normal	Limited	
N	10	11	10	
Activity score/min	144 (71–221)	120 (49–324)	50 (29–257)[a]	0.0199
Discrepancy%	0 (−6 to +5)	+6 (−1 to +21)*	+6 (−6 to +31)*	0.0161

Data from subjects with average activity score <26/min were rejected. Medians and ranges are given. Discrepancy was calculated as in Table 5; *different from Group I, $P<0.05$, Dunn's test.

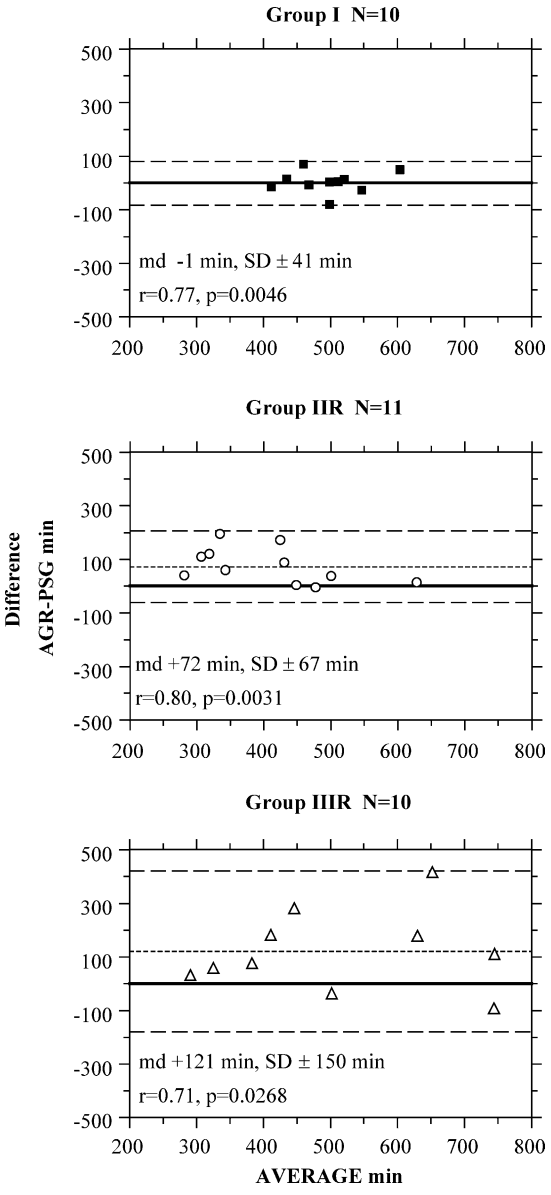

Fig. 2. Differences between actigraphic and polysomnographic total sleep estimation plotted against the average total sleep obtained with the two methods. The dotted line denotes the mean difference, broken lines the range of ±2 standard deviations of the differences. Data are from the reduced groups (IIR, IIIR) from which the recordings with an average activity score < 26/min were rejected.

In these selected recordings, total sleep estimated by AGR was significantly related to total sleep scored in PSG.

4. Discussion

In the group of normally moving subjects with normal sleep, the relatively simple sleep scoring algorithm of Actiwatch Sleep Analysis resulted in sleep parameters that agreed with PSG comparably to the parameters obtained by some other algorithms [16–18]. De Souza and collaborators [19] systematically compared Cole's and Sadeh's algorithms in a group of 21 healthy volunteers, using concordance calculations similar to those used in the present study [14]. There were no major differences between the results obtained by Cole's and Sadeh's algorithms: both overestimated total night sleep slightly (by 8–18 min), resulting in a 2–4%-overestimation of efficiency. Another algorithm, suggested by Jean-Louis and co-workers, [18] yielded total night sleep time 10 min shorter (calibration group) or 12 min longer (validation group) than that determined by PSG. In our healthy volunteers, the mean underestimation of night sleep by Actiwatch Sleep Analysis was 17 min, and efficiency assessed by AGR was 3% lower than that by PSG. In our analyses, the MT epochs in PSG were regarded as sleep, which in part explains the underestimation of sleep by AGR.

In an extensive recent study of 14 healthy volunteers, parallel AGR and PSG recordings were performed during seven consecutive 24 h periods [20]. Two different approaches were used to define AGR sleep and wake over the recordings. Based on calculations of AGR predictive values for PSG states, it was concluded that both methods overestimated sleep by 1.3/24 h. The authors point out that the discrepancy in part resulted from 'relative insomnia', because the volunteers in laboratory conditions often stayed in bed longer than they were able to sleep. This occurred in only two of our 10 volunteers. In addition, we used different thresholds for defining sleep for the periods in bed and out of bed. On average, the overestimation of daytime sleep compensated the underestimation of night sleep in our healthy volunteers.

In the patients with sleep disorders and normal motor abilities, sleep was highly overestimated by AGR, even after omitting the recordings of two almost motionless subjects. The rejected recordings were from patients on heavy antiepileptic and psychiatric medication; both individuals were somnolent during daytime and one was diagnosed to have a severe sleep-related breathing disorder. In a previous study [6], the same scoring algorithm was applied to overnight recordings from 100 sleep-disordered patients; the mean overestimation of total sleep time (1.0 or 1.4 h; low or medium threshold, respectively) is very similar to our 60 min overestimation of night sleep in the original Group II. The overestimation of sleep has also been a common problem in applying other types of devices and scoring methods to some groups of patients suffering from disturbed sleep [2,3], and optimizations of the scoring methods have been suggested for special patient groups [4,21]. In our patients, sleep onset insomnia seemed to be the main source of discrepancy between AGR and PSG. This suggests that a redefinition of sleep onset in Actiwatch Sleep Analysis might increase its agreement with PSG for sleep-disordered subjects. The objectivity of AGR becomes questionable, however, if we select a scoring method according to a diagnosis assessed beforehand.

In the group of subjects with sleep disorders and motor handicaps, there was an even greater overestimation of sleep

by AGR. In the original group of unselected subjects, most sleep parameters assessed by the two methods were not even related. The most obvious reason for the discrepancy is that a method based on acceleration measurements and algorithms, developed for normally moving people, cannot function equally well in subjects with abnormal or limited movements—or in motionless persons. Indeed, some practically motionless subjects were originally included in Group III. Our idea was to show that, by eliminating the subjects with the most severe motor handicaps, an acceptable agreement between AGR and PSG sleep measures might be found. The aim was to determine 'the lowest motor ability score' for reliable AGR sleep estimation.

To our surprise, the clinical evaluation of motor abilities was not a good predictor for agreement. There were individuals whose neurological examination disclosed rather slight limitations in moving, but in practice did not use their abilities, and a few handicapped individuals, clinically predicted to be poor candidates for AGR sleep assessment, for whom the sleep parameters determined by the two methods were well in line. This applied especially to the subjects who did not move voluntarily but frequently displayed waking stereotypic arm movements similar enough to normal arm movements to bypass the accelerometer signal filtering.

The actual activity of the subjects, defined as average activity score/min during the recording, was a better predictor to the agreement between AGR and PSG than clinical evaluation of motor abilities. For subjects with an average activity score >25/min, the total sleep time assessed by AGR was related to that assessed by visual scoring of PSGs with the device and standard algorithms used in this study. However, the marked overestimation of sleep by actigraphy remained. We conclude that standard AGR can provide some information about the sleep times of subjects with motor handicaps if the patient's actual activity exceeds a certain limit. The sleep parameters may be related, but not equal, to those assessed by PSG. Because different types of devices produce different 'activity counts', our threshold score cannot be applied to other devices.

In addition to the motor disability, there are other factors that might have worsened the agreement between AGR and PSG sleep estimation, especially in Group III. First, the sleep disorders in the original Group III were more severe than in Group II (Tables 1 and 3). We have no direct information about the contribution of disordered sleep to the poor agreement of the sleep measures in Group III. Locomotor disability is highly associated with sleep disorders in mentally retarded people [22], and a group of subjects with motor handicaps and normal sleep was not available for comparisons. The elimination of the recordings with very low activity counts caused a concomitant exclusion of the most abnormal sleep patterns from the analyses. The discrepancy% between AGR and PSG did not differ significantly between Groups IIR and IIIR, suggesting that the discrepancy among the selected Group IIIR subjects might at least partly originate from their disordered sleep and not from their limited motor abilities.

Secondly, the discrepancy between AGR and PSG sleep estimation in subjects with developmental brain disorders may also arise from difficulties in defining sleep in PSG [23], a problem similarly occurring in studies of elderly people with degenerative brain diseases [12]. Several subjects had EEG abnormalities causing difficulties in sleep scoring, especially in assessing stage transitions, which occurred frequently in these subjects. Several subjects in Group II had healthy electroencephalograms, while all subjects in Group III had developmental brain disorders.

Although the sensitivity of the actigraphs used in this study is high enough to give reliable results when applied to normally moving people, more sensitive devices might improve the differentiation of sleep and wake when recording the handicapped. There is also some variation in the sensitivity of the devices themselves, either at the outset or over the course of time. Such differences did not significantly affect our results because six watches were used randomly in each group.

A fourth factor that may have caused deterioration in the agreement between AGR and PSG in all three groups was the PSG recording itself; the equipment, although ambulatory, limited the subjects' freedom to move. Additional 7-day AGR recordings (with the same device as used in the study) were available from most of our subjects. In these recordings the average activity scores were usually clearly higher than during the PSG recording. The electrodes, wires and portable battery box reduced the activity of the healthy volunteers, on average to 60% of that on normal days. In most patients, the activity counts also fell, in some of them to 30% of normal. A few patients were probably agitated by the study and moved more than usual. Since PSG is used as a standard method for comparisons, its possible effects on the measurements cannot be avoided.

The night sleep scoring software used in this study enables the choice of three levels of differentiating threshold for wake and sleep (activity score 20, 40 or 80/min). As far as we know, there are no studies providing guidelines for this selection. Originally, we performed all analyses with all three thresholds and chose the threshold that gave the best agreement with PSG. It appeared that the limit in the application of low and medium thresholds was an average activity score of 100/min. High threshold was not usable in our subjects. The Nap Analysis software is even more flexible, requiring the definition of activity level and duration of immobility for 'sleep'. We tested many combinations and found many individual 'best thresholds' ($=$the highest agreement with PSG). For unity, the results are given with two thresholds, depending on the average activity count/min used in the night sleep scoring. There is clearly a need for systematic studies to achieve rules for threshold selection, both in Actiwatch Sleep Analysis

and Nap Analysis software. In general, the use of different scoring algorithms for the periods in bed and out of bed may be justified, because sleep is more probable in bed than out of bed. Others have also come to this conclusion [21].

Epoch-by-epoch comparisons were not performed between AGR and PSG due to different epoch lengths, insufficient synchronization of the timers, and because the marked overestimation of sleep showed that the sensitivity of AGR to detect PSG sleep was probably high and the ability to detect wake (specificity) was low. The accuracy could not be higher than that calculated by Kushida and collaborators [6], who used the same algorithm to analyse recordings of 100 sleep-disordered patients (0.77 ± 0.11 with low threshold for night sleep). Finally, it is not possible that arm movements can measure exactly the same physiological states reflected in multiple electrophysiological measures.

In summary, the clinical evaluation of motor disability was not as useful as the measurement of overall, actual activity in predicting the agreement between actigraphic and polysomnographic measures of sleep. Excluding the recordings of patients with very low activity (due to motor handicap or other reasons), significant correlation was found between AGR and PSG in total sleep estimation. The inspection of the actigram itself can be used to evaluate its usability in sleep estimation of handicapped patients. The absolute values of sleep parameters calculated from AGR recordings can be quite different from those measured by PSG In this study, overestimation of sleep by ARG was the most common source of discrepancy.

Acknowledgements

This work was financially supported by the Academy of Finland and Rinnekoti Research Foundation, Espoo, Finland. We thank Ms Aliisa Sinivaara and Ms Kirsi Koistinen for technical assistance.

References

[1] Sadeh A, Hauri PJ, Kripke DF, Lavie P. The role of actigraphy in the evaluation of sleep disorders. Sleep 1995;18:288–302.
[2] Sadeh A, Acebo C. The role of actigraphy in sleep medicine. Sleep Med Rev 2002;6:113–24.
[3] Ancoli-Israel S, Cole R, Alessi C, et al. The role of actigraphy in the study of sleep and circadian rhythms. Sleep 2003;26:342–92.
[4] Jean-Louis G, Mendlowicz MV, Gillin JC, et al. Sleep estimation from wrist activity in patients with major depression. Physiol Behav 2000;70:49–53.
[5] Hauri PJ, Wisbey J. Wrist actigraphy in insomnia. Sleep 1992;15: 293–301.
[6] Kushida CA, Chang A, Gadkary C, et al. Comparison of actigraphic, polysomnographic, and subjective assessment of sleep parameters in sleep-disordered patients. Sleep Med 2001;2:389–96.
[7] Heikkilä E, Hätönen TH, Telakivi T, et al. Circadian rhythm studies in neuronal ceroid-lipofuscinosis (NCL). Am J Med Gen 1995;57: 229–34.
[8] Van Someren EJ, Kessler A, Mirmiran M, Swaab DF. Indirect bright light improves circadian rest-activity rhythm disturbances in demented patients. Biol Psychiatry 1997;41:955–63.
[9] Mishima K, Okawa M, Satoh K, et al. Different manifestations of circadian rhythms in senile dementia of Alzheimer's type and multi-infarct dementia. Neurobiol Aging 1997;18:105–9.
[10] Hätönen T, Kirveskari E, Heiskala H, et al. Melatonin ineffective in neuronal ceroid-lipofuscinoses patients with fragmented or normal motor activity rhythm recorded by wrist actigraphy. Mol Gen Metab 1999;66:401–6.
[11] Motohashi Y, Maeda A, Wakamatsu H, et al. Circadian rhythm abnormalities of wrist activity of institutionalized dependent elderly persons with dementia. J Gerontol 2000;55A:M740–M743.
[12] Ancoli-Israel S, Clopton P, Klauber MR, et al. Use of wrist activity for monitoring sleep/wake in demented nursing-home patients. Sleep 1997;20:24–7.
[13] Carskadon MA, Rechtschaffen A. Monitoring and staging human sleep. In: Kryger MH, Roth T, Dement WC, editors. Principles and practice of sleep medicine, 3rd ed. Philadelphia: Saunders; 2000, p. 1197–215.
[14] Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurements. Lancet 1986;8: 307–10.
[15] Conover WJ. Practical nonparametric statistics. New York: Wiley; 1971.
[16] Sadeh A, Alster J, Urbach D, Lavie P. Actigraphically based automatic bedtime sleep–wake scoring: validity and clinical applications. J Ambul Monit 1989;2:209–16.
[17] Cole RJ, Kripke DF, Gruen W, et al. Automatic sleep/wake identification from wrist activity. Sleep 1992;15:461–9.
[18] Jean-Louis G, von Gizycki H, Zizi F, et al. Determination of sleep and wakefulness with the actigraph data analysis software (ADAS). Sleep 1996;19:739–43.
[19] de Souza L, Benedito-Silva AA, Nogueira Pires ML, et al. Further validation of actigraphy for sleep studies. Sleep 2003;26:81–5.
[20] Pollak CP, Tryon WW, Nagaraja H, Dzwonczk R. How accurately does wrist actigraphy identify the states of sleep and wakefulness?. Sleep 2001;24:957–65.
[21] Jean-Louis G, Kripke DF, Cole RJ, et al. Sleep detection with an accelerometer actigraph: comparisons with polysomnography. Physiol Behav 2001;72:21–8.
[22] Lindblom N, Heiskala H, Kaski M, et al. Neurological impairments and sleep–wake behaviour among the mentally retarded. J Sleep Res 2001;10:309–18.
[23] Espie CA, Paul A, McFie J, et al. Sleep studies of adults with severe or profound mental retardation and epilepsy. Am J Ment Retard 1998; 103:47–59.