Actigraphy combined with EEG compared to polysomnography in sleep apnea patients

Ingo Fietze¹, Thomas Penzel¹, Markku Partinen², Jochen Sauter³, Gert Küchler⁴, Alexander Suvoro⁵ and Holger Hein⁶

¹ Sleep Medicine Center, Charité Universitätsmedizin Berlin, Germany  
² Helsinki Sleep Clinic and Department of Neurology, University of Helsinki, Helsinki, Finland  
³ Sleep Laboratory, Freiburg, Germany  
⁴ SOMNOmedics GmbH, Randersacker, Germany  
⁵ Institute Medico-biological Problems, Moscow  
⁶ Sleep Laboratory, Reinbeck, Germany

E-mail: thomas.penzel@charite.de

Received 11 October 2014, revised 17 December 2014  
Accepted for publication 30 December 2014  
Published 5 February 2015

Abstract

An actigraph extended with electroencephalography (EEG), electroocculography (EOG) and electromyography (EMG) was compared to polysomnography in two studies on patients suffering from sleep disordered breathing. Study A with 30 subjects used a single lead EEG, and study B with 20 subjects used EOG and EMG in addition. Sleep was scored according to Rechtschaffen and Kales rules. Total sleep time (TST), sleep period time (SPT), sleep efficiency (SE), sustained sleep efficiency (SSE), sleep-onset latency (SL), and sleep stages were compared. For study A an epoch-by-epoch comparison of sleep stages revealed an overall agreement of 74.2%. Correlations were high for SE (0.98, p < 0.001), SSE (0.98, p < 0.001), TST (0.99, p < 0.001), SPT (0.99, p < 0.001), and SL (0.98, p < 0.001). Regarding the sleep stages, correlations were high for rapid eye movement (REM) (0.83, p < 0.001), light-sleep (0.78, p < 0.001), and deep sleep (0.73, p < 0.001). For study B, results of an epoch-by-epoch comparison of sleep stages showed an overall agreement of 75.5%. Correlations were high for SE (0.98, p < 0.001), SSE (0.98, p < 0.001), TST (0.87, p < 0.001), SL (0.98, p < 0.001), SSE (0.94, p < 0.001), and for rapid eye movement (REM) (0.91, p < 0.001), light-sleep (0.74, p < 0.001), and deep sleep (0.89, p < 0.001). In summary the study revealed high agreement between polysomnography and single lead EEG in sleep apnea patients. Deviations for REM were slightly higher for the single
lead EEG compared to single lead EEG plus EOG/EMG. Both simplified systems proved to be reliable for comfortable out-patient sleep recording.

Keywords: polysomnography, validation, EEG, sleep apnea, sleep stages

(Some figures may appear in colour only in the online journal)

1. Introduction

Unattended sleep recording is not only possible but indeed enables good-quality results, even for polysomnography (PSG) (Redline et al 1998). Single lead EEG recording had been tested as a simple method for home sleep monitoring. Stepnowsky et al (2013) demonstrated the accuracy of visual sleep-stage scoring using a single lead EEG recording device. Satisfactory analysis of a single channel EEG is also possible (Koley and Dey 2012, Kosmadopoulos et al 2014). As early as 2004, Fischer et al showed that a visually evaluated single channel EEG is effectively suited for characterizing sleep, and also for snorers and patients with obstructive sleep apnea (OSA). Gfüllner and Siemon (2000) had previously raised concerns on this reduced recording practice. Actigraphy without additional signals is used as an indirect way to study sleep (Jean-Louis et al 2001, Weiss et al 2010, Shambroom et al 2012, Kosmadopoulos et al 2014). Actigraphy had been applied in patients with sleep-related breathing disorders (Dick et al 2010). However, this procedure is evidently more difficult for older persons (Siversten et al 2006). The study conducted by Stepnowsky et al (2013) also investigated OSA patients. The authors discussed a dependence of automatic single channel EEG measurement on the following factors: apnea-hypopnea index (AHI), medication, and signal quality. This signifies that the value of a single channel EEG in patients with OSA had not yet been clarified. The goal of the present study is to investigate the accuracy of a single lead EEG recorder in comparison to standard PSG, as well as a single lead EEG plus two-channel EOG and EMG compared to PSG.

2. Methods

We performed 30 recordings (study A) and 20 recordings (study B) in a sleep laboratory, with simultaneous recording of PSG and actigraphy plus single channel EEG (F4-M1) (study A), as well as PSG and actigraphy plus single channel EEG (F4-M1) plus two-channel EOG and EMG (chin) (modified R&K EEG) (study B).

2.1. Participants

For study A subjects had an average age of 57 ± 14 years (5 female, 25 male), with a range from 18 to 80 years. The average body mass index (BMI) was 31 ± 5 kg m\(^{-2}\) with a range from 23 to 45 kg m\(^{-2}\). For all subjects the diagnosis of obstructive sleep apnea syndrome (OSAS) was confirmed. The polysomnographic recordings were performed during the diagnostic night (11 recordings) or during the positive airway pressure (PAP) treatment night (19 recordings), with 18 patients on continuous positive airway pressure (CPAP) and one on bi-level positive airway pressure (BiPAP) therapy. In six subjects, a periodic leg movement syndrome (PLMS) was diagnosed, and one subject suffered from arrhythmia absoluta.

In the second approach (study B), subjects had an average age of 60 ± 11 years (6 female, 14 male), with a range from 41 to 74 years. The average BMI was 31 ± 6 kg m\(^{-2}\), with a range from 22 to 40 kg m\(^{-2}\). For all subjects the diagnosis of OSAS was confirmed. Polysomnographic
recordings were conducted during the diagnostic night (six recordings) and therapy night (14 recordings), with 11 patients on CPAP, one on BiPAP, and two patients wearing a supine position prevention vest. In 12 subjects, PLMS was found and Cheyne–Stokes respiration was noted in one subject.

2.2. Polysomnography

The Somnoscreen system (SSC; Somnomedics, Germany) was applied as an online wireless PSG device. PSG was configured to record EEG leads C4, C3, A2, A1, EOG left eye, EOG right eye, EMG, ECG, flow (cannula), snoring sounds, respiratory effort signals, SpO2, pulse rate, finger plethysmogram, ambient light, and body position. PSG data were manually scored using a consensus scoring from three certified sleep practitioners in random order. Sleep staging was conducted by applying the standardized Rechtschaffen and Kales (R&K) rules with 30 s epoch windows. The hypnogram from PSG was used as a reference result for all comparisons. PSG recordings were performed online because the Somnoscreen is calibrated with computer time only in the online mode. Sleep statistics derived from the hypnogram were obtained from the sleep report provided by the proprietary Domino software (table 1), which was used for all three hardware settings. In addition, time in bed (TIB) was determined from the PSG (Rodenback 2013).

Stages one and two were summarized to light sleep and stages 3 and 4 were summarized to deep sleep. Additionally rapid eye movement (REM) sleep was scored.

2.3. Single channel EEG (study A)

The Somnowatch (SOW; Somnomedics, Germany) records the acceleration on three axes and is calibrated in mG. The x, y, and z axes can be stored separately or as magnitude signals, corresponding to the square root of the sum of each of the three axes squared. Sensitivity extends to 0.004 G, with a range of ± 8.7 G. An external channel allows recording of an additional biosignal and served to record an EEG (F4-M1) with a sampling rate of 256 Hz. We applied the actigraphy recorder on the patients’ chests to obtain additional information on body position, which was recorded at 1 Hz. The light signal was stored every 30 s. After the recording, the raw data were transferred to a computer and analyzed automatically. The ‘lights on’ and

<table>
<thead>
<tr>
<th>Table 1. Definitions of sleep parameters for the DOMINO light software.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
</tr>
<tr>
<td>TIB (min)</td>
</tr>
<tr>
<td>TST (min)</td>
</tr>
<tr>
<td>SE (%)</td>
</tr>
<tr>
<td>SL (min)</td>
</tr>
<tr>
<td>SSE (%)</td>
</tr>
<tr>
<td>SPT (min)</td>
</tr>
<tr>
<td>REM (%)</td>
</tr>
<tr>
<td>LS (%)</td>
</tr>
<tr>
<td>DS (%)</td>
</tr>
</tbody>
</table>

TIB = time in bed; TST = total sleep time; SE = sleep efficiency; SL = sleep onset latency; SSE = sustained sleep efficiency; SPT = sleep period time; REM = rapid eye movement sleep; LS = light sleep; DS = deep sleep.
‘lights off’ markers were set manually in SOW data to match PSG data, which were mostly of varying lengths. Single channel EEG data were manually scored using a consensus scoring from three certified sleep practitioners in random order. Sleep staging for single channel EEG was performed visually with 30 s epoch windows, according to R&K, but not using the criteria for EOG and EMG because these signals were not available. REM sleep was scored if eye movements were seen in the frontal lead EEG. As an additional criterion for REM sleep, sigma wave activity was checked in the EEG and should not have been present. The absence of sleep spindles was also required during REM sleep.

2.4. Single channel EEG plus EOG plus EMG (modified R&K EEG) (study B)

For the second approach (study B), the external channel of the Somnowatch was used to record four additional biosignals: EEG (F4-M1) with a sampling rate of 256 Hz, EMG (chin) with a sampling rate of 256 Hz, as well as EOG left eye and EOG right eye, both with a sampling rate of 128 Hz. The device was applied at the patient’s chest. The light signal and the body position were stored every 30 s. Sleep staging was conducted by applying standardized R&K rules, with 30 s epoch windows and data were manually scored analogous to PSG and single channel EEG data.

2.5. Statistical analysis

We calculated Pearson correlations and plotted Bland–Altman plots (Bland and Altman 1986). Because correlation assures only relative and not absolute validity, we also included an epoch-by-epoch comparison to check for sensitivity (i.e. how effectively light-sleep and slow-wave sleep are detected), specificity (how effectively REM sleep is detected), and overall accuracy or agreement. Sensitivity was defined as the ability of the single lead EEG to detect light sleep (stage 1 and 2) or slow-wave sleep when PSG showed the corresponding sleep stage. Specificity was defined as the ability to detect waking and REM when in PSG, which was scored as ‘wake’ and ‘REM’. The epoch-by-epoch analysis was performed with Matlab and the validity of the results was confirmed by using a test file with a predefined sleep-stage percentage and error ratio.

3. Results

3.1. Single channel EEG versus PSG (study A)

TIB was identical for the PSG and the single channel EEG device, because the markers that define TIB in single channel EEG data were transferred from the PSG data to SOW data without changes. Comparison of sleep data are presented in table 2. Analysis revealed a strong positive correlation of $r = 0.99$ between PSG and single channel EEG on total sleep time (TST) and on sleep period time (SPT), as well as between sleep efficiency (SE), sustained sleep efficiency (SSE), and sleep-onset latency (SL) in PSG and single channel EEG. A significant correlation was also found for light and deep sleep stages and for REM sleep. However $r$ values were somewhat lower than the macrosleep parameters (table 2).

The Bland–Altman plot for TST (figure 1) shows that the PSG scores are on average 10 min higher than the single lead sleep EEG scores (10 ± 12 min). This signifies that the single lead EEG Somnowatch underestimated TST on average by approx. 10 min. The differences are likely to be normally distributed and thus it is fairly safe to assume that 95% of all measurement differences for TST will lie between the agreement lines of −35 and 15 min with the exception for one value lying beyond this area.
The Bland–Altman plot shows that for SPT the PSG scores are on average $3 \pm 4$ min higher than the single lead EEG score (figure 2). As shown for TST, the differences for SPT are also likely to be normally distributed. This likewise indicates that 95% of all measurement differences for SPT will lie between the agreement lines of $-11$ and $5$ min.

The Bland–Altman plot shows that, for SE, PSG scores are on average $3 \pm 4\%$ higher than the single lead EEG score (figure 3). The differences are likely to be normally distributed, which again allows the safe assumption that 95% of all measurement differences for SE will lie between the agreement lines of $-4$ and $10\%$.

The Bland–Altman plot shows that, for SSE, PSG scores are on average $2 \pm 4\%$ higher than the actigraphy score (figure 4). The differences are likely to be normally distributed, which allows the safe assumption that 95% of all measurement differences for SSE will lie between the agreement lines of $-9$ and $5\%$.

The Bland–Altman plot shows that for SL the polysomnographic scores are on average $3 \pm 6$ min lower than the single lead EEG scores (figure 5). As before, the differences are likely to be normally distributed, which allows the fairly safe assumption that 95% of all measurement differences for sleep period times will lie between the agreement lines of $-10$ and $16$ min.
3.2. Overall accuracy

The epoch-by-epoch comparison between PSG and single lead EEG, with the PSG data as standard, showed that 74.2% of the sleep epochs were correctly identified. Light sleep during TST from one-channel EEG and PSG correlated with \( r = 0.71 \), whereas one-channel EEG overestimated light sleep by 2% (ranging from 39% overestimation to 19% underestimation). Deep sleep during TST from one-channel EEG and PSG correlated with \( r = 0.73 \), whereas one-channel EEG underestimated mean slow-wave sleep by 1% (ranging from 30% overestimation to 18% underestimation). REM sleep was likewise effectively detected, with a correlation of \( r = 0.83 \) and with underestimation of REM during TST from one-channel EEG by 2% (ranging from 4% overestimation to 16% underestimation).
3.3. Single channel EEG plus EOG plus EMG /modified R&K EEG (study B)

As for study A, TIB was identical for both devices, since the markers that define TIB in modified R&K-EEG data were transferred from SSC data to SOW data without changes. On average, modified R&K-EEG underestimated TST by $23 \pm 27$ min, SPT by $2 \pm 11$ min, SE by $5 \pm 7\%$, and SSE by $4 \pm 7\%$.

3.4. Comparison of sleeping times

Analysis revealed a positive correlation of $r = 0.87$ between PSG and modified R&K-EEG for TST, $r = 0.94$ for SPT, $r = 0.86$ for SE, $r = 0.98$ for SSE, and $r = 0.98$ for SL (table 3). TIB for modified R&K EEG data were transferred from the PSG data without changes. A significant
correlation was also found for light and deep sleep stages and for REM sleep. This was specifically true for deep sleep \((r = 0.89)\) and REM sleep \((r = 0.91)\).

The Bland–Altman plot for TST shows that the polysomnographic scores are an average of 23 min higher than the modified R&K EEG score (figure 6). The differences are likely to be normally distributed, which enables the fairly safe conclusion that all measurement differences for TST will lie between the agreement lines of 29 and −75 min.

The Bland–Altman plot shows that for SPT the PSG scores are on average 2 ± 11 min higher than the modified R&K EEG score (figure 7). The differences are likely to be normally distributed, and it is therefore fairly safe to say that 90% of all measurement differences for SPT will lie between the agreement lines of 20 and −24 min. Only two values lie beyond this area.

The Bland–Altman plot shows that for SE, the PSG scores are an average of 5 ± 6% higher than the modified R&K EEG score (figure 8). The differences are likely to be normally distributed, which allows the fairly safe assumption that 95% of all measurement differences for SE will lie between the agreement lines of 8 and −18%.

The Bland–Altman plot shows that for SSE the PSG scores are an average of 4 ± 7% higher than the modified R&K EEG score (figure 9). The differences are likely to be normally distributed, which means that it is fairly safe to say that 95% of all measurement differences for SSE will lie between the agreement lines of 9 and −17%.

The Bland–Altman plot shows that for SL the polysomnographic scores are an average of 4 min lower than the modified R&K EEG scores (figure 10). The differences are likely to be

---

Table 3. Pearson correlations for actigraphy versus PSG for study B.

<table>
<thead>
<tr>
<th></th>
<th>TST</th>
<th>SPT</th>
<th>SE</th>
<th>SSE</th>
<th>SL</th>
<th>TIB</th>
<th>DS</th>
<th>LS</th>
<th>REM</th>
</tr>
</thead>
<tbody>
<tr>
<td>(r_{Pearson})</td>
<td>0.87</td>
<td>0.94</td>
<td>0.86</td>
<td>0.86</td>
<td>0.87</td>
<td>1.00</td>
<td>0.89</td>
<td>0.74</td>
<td>0.91</td>
</tr>
<tr>
<td>(p)-value</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Mean R&amp;K-EEG</td>
<td>324 min</td>
<td>386 min</td>
<td>79.4%</td>
<td>82.7%</td>
<td>21 min</td>
<td>408 min</td>
<td>23.2%</td>
<td>55.1%</td>
<td>21.8%</td>
</tr>
<tr>
<td>Mean PSG</td>
<td>347 min</td>
<td>388 min</td>
<td>84.9%</td>
<td>86.8%</td>
<td>17 min</td>
<td>408 min</td>
<td>20.7%</td>
<td>59.5%</td>
<td>18.6%</td>
</tr>
</tbody>
</table>

TST = total sleep time; SPT = sleep period time; SE = sleep efficiency; SSE = sustained sleep efficiency; SL = sleep onset latency; TIB = time in bed; DS = deep sleep; LS = light sleep; REM = rapid eye movement sleep.

---

Figure 6. Bland Altman TST. Average R&K EEG from SOW and PSG from SSC versus difference in TST of both methods. There is a shift of −23 min, which means that the SOW underestimated TST by about 23 min on average.
normally distributed, and it is fairly safe to say that 95% of all measurement differences for SPT will lie between the agreement lines of 20 and −12 min.

3.5. Overall accuracy

The epoch-by-epoch comparison between PSG and modified R&K PSG, with PSG data as the standard, showed that 75.5% of the sleep epochs were correctly identified. Light sleep during TST from modified R&K EEG and PSG correlated with $r = 0.74$, while single lead EEG underestimated light sleep by a mean of 4% (ranging from 18% overestimation to 22% underestimation). Slow-wave sleep during TST from modified R&K EEG and PSG correlated

![Figure 7. Bland Altman SPT. Average R&K EEG from SOW and PSG from SSC versus difference in SPT of both methods. There is a shift of ~2 min, which means that the SOW underestimated SPT by about 2 min on average.](image)

![Figure 8. Bland Altman sleep efficiency (SE). Average R&K EEG from SOW and PSG from SSC versus difference in SE of both methods. There is a shift of ~5%, which means that the SOW underestimated SE at a level of 5% on average.](image)
with $r = 0.8$, while one-channel EEG overestimated deep sleep by a mean of 3% (ranging from 13% overestimation to 7% underestimation). REM sleep was detected with a correlation of $r = 0.91$ overestimating REM during TST from single lead EEG by 3% (ranging from 10% overestimation to 4% underestimation).

4. Discussion

This study determined good agreement among OSA patients in standard sleep parameters between PSG and single lead EEG or modified R&K EEG. Agreement was slightly better for the modified R&K EEG system than for the single lead EEG system, which attests for the validity of visual analysis (with analysis improving with additional recording channels). The new SOMNOwatch™ EEG system (SOMNOmedics, Germany) has thereby proved to be a reliable tool for out-patient sleep analysis, including OSA patients, whether one single lead EEG or EEG in combination with EOG and EMG is applied.
Although the agreement between single lead EEG and PSG was lower than the R&K module compared to PSG, as expected, both methods are equally effective to determine macro-sleep parameters like TST, SE and sleep latency. There is even a significant correlation in the detection of REM and light and deep sleep stages using the single lead EEG in comparison to PSG in epoch-by-epoch analysis. Special attention was directed to the lack of spindles and K-complexes as well as to eye movements reflected in the frontal single lead EEG, which explains the accuracy.

The study accordingly contributes to the ability to effectively measure sleep variables, within the context of the increase in polygraphic examinations, with a minimum of expense and effort that may be required. Single lead EEG suffices to effectively evaluate sleep structure, SE, and length of sleep for OSA patients.

Also focusing on patients with OSA, but performing an automatic evaluation, Stepnowsky et al. (2013) achieved a high degree of correlation to PSG, both by evaluation according to American Academy for Sleep Medicine (AASM) and to R&K criteria. In this context, the degree of OSA, medications, signal quality, and the experience of the raters exert influence on analysis of the frontal EEG.

Koley and Dey (2012) likewise established a correlation higher than $r = 0.85$ between automatic single channel evaluation and PSG. Kosmadopoulos et al. (2014) investigated healthy subjects with single channel EEG and automatic evaluation and compared results to PSG. They determined an epoch agreement of 91.6% between single channel EEG and PSG for indentifying sleep and wake. Using the same EEG headband system, Griessenberger et al. (2013) found underestimation of the waking state, in comparison to visually as well as automatically analyzed PSG. Shambroom et al. (2012) reported epoch correlation of 81.1% between automatically analyzed headband EEG and visually evaluated PSG.

In 2004 Fischer et al. investigated the QUISI system and, on the basis of visual analysis, determined sensitivity and specificity of over 0.9 in applications with sleep apnea patients. Their study revealed, with respect to the individual stages of sleep, underestimation of TST, SE, stage 2, and slow-wave sleep (SWS), as well as overestimation for the waking phase, all within a range of approximately 5%.

The correlations determined in our study between the parameters of the macrostructure of sleep (TST, SPT, SE, SL) with the standard PSG were high and, independently of the device, not less than $r = 0.85$. With regard to sleep microstructure—and to the distribution of the light-sleep stages (1 and 2), SWS, and REM sleep—we similarly determined a high degree of correlation within the range of $r = 0.7$–0.9, with the best correlation for REM sleep. The latter is above all noteworthy for the single channel EEG system, and is most heavily influenced by scorer experience.

High deviation, including high variability (standard deviation), was apparent especially for TST, as recorded with the R&K Somnowatch device. This may be explained by differences in the patient cohort and also by possible quality differences in application of the measuring systems. Groups A and B were recorded in two different sleep laboratories which might explain further differences.

Precisely this aspect must be critically noted: i.e. that we investigated the accuracy of two different device configurations in comparison to the standard PSG for two different patient cohorts. As a result we can reach no conclusion on the comparison of the two outpatient configurations: single lead EEG versus modified R&K EEG. We also did not study possible influencing factors for validity of the single lead EEG recording: e.g. the influence of respiratory disturbances, the therapy of these disturbances with CPAP, or concomitant diseases.

Nevertheless, we conclude that outpatient single lead EEG recording with or without EOG and EMG appears effective, even for OSA patients, e.g. for objectifying deficits in deep sleep
and/or dream sleep as well as for SE. The precondition here is manual evaluation of the data, whereas in the near future automatic analysis and new analytical procedures will grow in importance (see García Correa et al 2014, Intiaz and Rodríguez-Villegas 2014).

Acknowledgment

Dr Anna Hennig and Dr Jennifer Gehring, Somnomedics, who prepared the statistics, the figures and parts of the text. Somnomedics provided Sensors and equipment at reduced costs.

References

Dick R, Penzel T, Fietze I, Partinen M, Hein H and Schulz J 2010 AASM standards of practice compliant validation of actigraphic sleep analysis from SOMNOwatch™ versus polysomnographic sleep diagnostics shows high conformity also among subjects with sleep disordered breathing Physiol. Meas. 31 1623–33
Grüllner F and Siemon G 2000 Studies with the fully automated EEG sleep analysis system QUISI Pneumologie 54 580–3
Shambrook J R, Fábregas S E and Johnstone J 2012 Validation of an automated wireless system to monitor sleep in healthy adults J. Sleep Res. 21 221–30
Stepnowsky C, Levendowski D, Popovic D, Ayappa I and Rapoport D M 2013 Scoring accuracy of automated sleep staging from a bipolar electrooculographic recording compared to manual scoring by multiple raters Sleep Med. 14 1199–207