Short sleep—poor sleep? A polysomnographic study in a large population-based sample of women

Torbjörn Åkerstedt1,2 | Johanna Schwarz1,2 | Georg Gruber3 | Jenny Theorell-Haglöw4 | Eva Lindberg4

1Department of Clinical Neuroscience, Karolinska Institute, Stockholm, Sweden
2Stress Research Institute, Stockholm University, Stockholm, Sweden
3The Siesta Group, Vienna, Austria
4Department of Medical Sciences, Respiratory, Allergy and Sleep Research, Uppsala University, Uppsala, Sweden

Correspondence
Torbjörn Åkerstedt, Department of Clinical Neuroscience, Karolinska Institute, Stockholm, Sweden. Email: torbjorn.akerstedt@ki.se

Funding information
The study was supported by The Swedish Research Council for Health, Working Life and Welfare, and by the Swedish Heart Lung Foundation.

Abstract
There is a lack of studies on the association between total sleep time (TST) and other polysomnographical parameters. A key question is whether a short sleep is an expression of habitual short sleep, or whether it reflects temporary impairment. The purpose of the present study was to investigate the association between TST and amount of sleep stages and sleep continuity measures, in a large population-based sample of women (n = 385), sleeping at home in a normal daily life setting. The results show that sleep efficiency, N1 (min), N2 (min), REM (min), REM% and proportion of long sleep segments, increased with increasing TST, whereas the number of awakenings/hr, the number of arousals/hr, N1% and REM intensity decreased. In addition, longer sleep was more associated with TST being perceived as of “usual” duration and with better subjective sleep quality. TST was not associated with habitual reported sleep duration. It was concluded that PSG-recorded sleep duration does not reflect habitual reported sleep duration in the present real-life context.

KEYWORDS
delta dominance, REM density, REM intensity, sleep spindles

1 | INTRODUCTION

Both long and short sleep have been associated with increased mortality and a number of other health problems (Cappuccio, Cooper, D’elia, Strazzullo, & Miller, 2011; Cappuccio, D’elia, Strazzullo, & Miller, 2010a, 2010b; Leng et al., 2015). Short and long sleep are also associated with obesity (Theorell-Haglow, Berne, Janson, Sahlin, & Lindberg, 2010; Theorell-Haglow & Lindberg, 2016).

Despite the apparent importance of sleep duration there have been very few studies describing the polysomnographical content of short and long sleep. In one study it was found that long sleepers with >8.5 hr of sleep had more REM sleep, but did not differ from short sleepers with respect to minutes of slow wave sleep (SWS) (Webb & Agnew, 1970; Webb & Friel, 1970), although the published tables suggest that the percentage of SWS should have been higher among the short sleepers. It was suggested that the increased REM may be a result of the dependence of REM sleep on sleep duration. Benoît, Foret, and Bouard (1983) confirmed the previous findings; so did Aeschbach, Cajochen, Landolt, and Borbély (1996), who found equal amounts of SWS (in minutes) in short and long sleepers, but also demonstrated a significantly higher percentage in short sleepers.
REM sleep was longer in long sleepers and also higher in percent. No other studies seem to be available on the polysomnography of habitual short and long sleep.

The studies of sleep duration discussed above were set in a laboratory environment, isolated from the influences of daily life, and optimized to promote sleep. These optimal sleep conditions may not be entirely representative of real-life sleep. In addition, the participants were students and in good health. This raises the question of what sleep would look like in a representative sample, on a normal night, recorded at home, in a setting of daily life chores before and after the recorded sleep. Would long and short sleep differ in sleep continuity or sleep stage content? This is quite a different question from that of how habitual long and short sleep differ. There is no previous work on this issue, but it might be the case that short sleep could indicate poor sleep in the sense of containing much superficial sleep and/or many awakenings. This hypothesis is based on our previous observation that low subjective sleep quality in the morning was closely related to short total sleep time (TST); low sleep efficiency and many awakenings (Akerstedt, Schwarz, Gruber, Lindberg, & Theorell-Haglow, 2016). We did not find any relation between sleep quality ratings and REM or SWS, and have difficulties formulating a definite hypothesis, but if short sleep indicates poor sleep, one may expect both SWS and REM to be reduced.

In addition, it might be of interest to investigate variables that represent sleep micro-architecture in relation to TST. Among them are arousals, power density in the delta band and other bands, sleep spindles and rapid eye movements (density and amplitude) (Feige et al., 2013). REM density is increased in depression, a disorder highly correlated with poor sleep in the sense of containing much superficial sleep and/or many awakenings. This hypothesis is based on our previous observation that low subjective sleep quality in the morning was closely related to short total sleep time (TST), low sleep efficiency and many awakenings (Akerstedt, Schwarz, Gruber, Lindberg, & Theorell-Haglow, 2016). We did not find any relation between sleep quality ratings and REM or SWS, and have difficulties formulating a definite hypothesis, but if short sleep indicates poor sleep, one may expect both SWS and REM to be reduced.

The purpose of the present study was to investigate the association between total sleep time (TST) and amount of sleep stages and sleep continuity measures, in a large cross-sectional sample of women, sleeping at home in a normal daily life setting. Women were selected for study because of their higher prevalence of sleep complaints.

2 | METHODS

2.1 | Design and participants

The study used a subsample of the Sleep and Health in Women (SHE) study (Sahlin, Franklin, Stenlund, & Lindberg, 2009). Within the SHE study as a whole, a representative sample of 10,000 women was randomly selected from the population registry of the City of Uppsala. These received a sleep/health questionnaire (response rate, 71.6%) and a second random sample of 400 (snorers were oversampled) was selected for polysomnography (PSG) recordings, (Theorell-Haglow, Berne, Janson, & Lindberg, 2008). Home PSG (including recording of obstructive sleep apnea [OSA] variables) was carried out for the 400 women. The mean age was 49.3 ± 11.2 (SD) years (22–72 years); the BMI was 25.3 ± 4.2, 73% were married/cohabiting, 51.7% had children at home, 44.2% had passed the menopause and 24.2% had received hormone treatment (ever). Fifteen participants turned out to have missing data on some central variables and thus 385 participants remained.

Treatment for sleep apnea (continuous positive airway pressure [CPAP] and oral appliances) was withdrawn for ≥3 nights preceding the PSG. This concerned only five cases. The original cohort has been reported on as a study of normative data on female sleep (Sahlin et al., 2009).

2.2 | Self-report variables

For information on subjective habitual sleep the Uppsala Sleep Questionnaire (Hetta, Broman, & Mallon, 1999) was used, with the following items: difficulties initiating sleep (DIS), difficulties maintaining sleep (DMS), early morning awakening (EMA), sleep quality, and tired daytime. The response was given on a scale from “none” to “very big problem” (1–5). Frequency of snoring was assessed by questions adopted from the Nordic Basic Sleep Questionnaire (Partinen & Gislason, 1995), with response options 1 (never) to 5 (every or almost every night). Daytime sleepiness was also assessed through the Epworth Sleepiness Scale (ESS) (Johns, 1991). The questionnaire also contained questions on the time of “lights out” and “time of awakening”. For anxiety and depression, the Hospital Anxiety and Depression Scale (HAD) (Zigmond & Snaith, 1983) was used. Depression and anxiety subscales ranged from 1 to 28. Use of hypnotics was rated on a scale from 1 to 5 (“never” to “most days”). Menopausal stage was rated through the question “Have you entered or passed menopause?”, with the response alternatives “yes”, “no” and “don’t know”. “Yes” and “don’t know” were combined in the analyses.

Upon awakening after the recorded sleep, participants rated: “Quality of sleep”, ranging from 0 to 100 (“good” to “bad”); “Sleep duration as usual” (0 to 100; “shorter” to “longer”); “Sleep quality as usual” (0–100; “better” to “worse”); and “Did anything happen during the day that may have disturbed sleep?” (0–1; “no” and “yes”).

2.3 | PSG recording

The PSG recording took place in the homes of the participants, using a solid-state portable sleep recorder (Embla, Flaga, Iceland). A standard electrode (silver/silver chloride) montage was used (C3, C4), referenced versus the contralateral mastoids, and in addition two submental electrodes and electrodes at the outer canthi of the eyes were used. To adapt to AASM scoring, F4 was interpolated. Further sensors included electrodes for bilateral anterior tibialis muscles, airflow with a three-port oronasal thermistor and a nasal flow pressure sensor, respiratory effort from piezo-electric belts (Resp-EZ, EPM Systems Midlothian, VA, USA), finger pulse oximetry (Embla A10 flex Sensor), electrocardiograms (VS), a piezo vibration sensor for snoring and a body position sensor. A research nurse applied the electrodes, connected the equipment, gave instructions in the early evening, and set the recording to start at an agreed-on time before the estimated
time of retiring. The equipment was retrieved the next morning by an experimenter. Data were lost for 1.5% (six of 400) of the participants but recording was repeated within a short period of time. Diary information was used to establish times for lights out and lights on.

Sleep staging and respiratory and arousal analysis were performed according to the classification criteria of the American Academy of Sleep Medicine (Iber, Ancoli-Israel, Chesson, Quan, & Medicine, 2007) using the computer-assisted sleep classification system Somnolyzer 24 x 7 (Anderer et al., 2005, 2010). All scoring was checked by a licensed sleep technician. Here, the terminology N1, N2 and N3 are used for sleep stages 1–3. Shift from any of the sleep stages to wake is expressed as awakenings per hour. Apnea was defined as a cessation of airflow for at least 10 s and hypopnea was defined as at least 10 s of 50% reduced nasal pressure, together with at least 3% desaturation. The apnea–hypopnea index (AHI) was defined as the mean number of apneas and hypopneas per hour of sleep. We also introduced a new variable to represent low frequency (0.5–4 Hz or delta) power. It was defined as the number of 30-s scoring epochs that contained ≥65% delta power (of total power in the 0.5–32 Hz band) in the epoch, expressed in percent TST. This variable was labelled delta dominance ≥65% (DD65%).

2.4 | Statistical analysis

Six categories of TST were formed: ≤5.5 hr (n = 74), 5.6–6.0 hr (n = 79), 6.1–6.5 hr (n = 71), 6.6–7.0 hr (n = 78), 7.1–8 hr (n = 53) and >8 hr (n = 23). The TST categories formed the fixed factor in a one-way analysis of variance, adjusted for age, BMI, somatic disease and use of medication (no adjusting for AHI). Trends across sleep durations were investigated through linear and quadratic contrasts. Analyses were carried out using SPSS, version 24. The significance level was set at p < 0.05. For some participants, the actual bedtime occurred somewhat before the recording started and these participants were excluded when computing the sleep latency, leaving n = 310 for that variable.

3 | RESULTS

The results show that sleep efficiency, N1 (min), N2 (min), REM (min), REM%, proportion of long sleep segments, delta dominance, N3 latency and REM latency increased with increasing TST, whereas the number of awakenings/hr, the number of arousals/hr, N1% and REM intensity decreased (Figures 1 and 2, Table 1). Of note is that N1 and N1% showed significant opposite trends, N1 increasing and N1% decreasing with increasing sleep duration. All significant F-

![Figure 1](image-url)
ratios had a linear trend ($p < 0.05$). Sleep efficiency, % sleep segments ≥ 10 minutes and % wake segments ≥ 10 minutes all had significant quadratic trends in addition to the linear trends.

In addition, long sleep was associated with higher-rated sleep quality, and ratings of relatively usual sleep duration and sleep quality (Figure 1 and Table 1). Long sleep was associated with early time for “lights out” and late times of rising. Having experienced something during the day that may have disturbed sleep (scored yes/no) showed a $\chi^2$ of 8.7 ($p = 0.123$) All significant F-ratios had a significant linear trend at $p < 0.002$, except for HAD depression (with $p = 0.011$). Quality of recovery sleep and sleep quality different from usual sleep also had significant quadratic trends ($p < 0.05$), in addition to the linear trends.

Habitual sleep quality (mean of three items) was not associated with TST, nor was HAD score for depression. HAD score for anxiety was lower in long TST, but this was almost entirely driven by the group of long sleepers (Table 1, Figure 1).

**4 | DISCUSSION**

The results show that sleep efficiency, N1 (min), N2 (min), REM (min), REM% and proportion of long sleep segments increased with
TABLE 1 Results from ANOVA of polysomnography (PSG) variables and self-reported data versus sleep duration. Adjusted for age, body mass index (BMI), disease and medication. df = 3/382, except for sleep latency (df = 3/307)

<table>
<thead>
<tr>
<th>Variables</th>
<th>F-ratio</th>
<th>p=</th>
<th>Variables</th>
<th>F-ratio</th>
<th>p=</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sleep efficiency</td>
<td>56.2</td>
<td>0.000*</td>
<td>Delta dominance</td>
<td>6.1</td>
<td>0.000</td>
</tr>
<tr>
<td>WTSP</td>
<td>22.7</td>
<td>0.000</td>
<td>REM density</td>
<td>1.3</td>
<td>0.263</td>
</tr>
<tr>
<td>Arousals/hr</td>
<td>7.1</td>
<td>0.000</td>
<td>REM intensity</td>
<td>8.2</td>
<td>0.000</td>
</tr>
<tr>
<td>% Sleep segments ≥ 10 min</td>
<td>4.2</td>
<td>0.001</td>
<td>% Wake segments ≥ 10 min</td>
<td>4.5</td>
<td>0.005*</td>
</tr>
<tr>
<td>N1%</td>
<td>1.1</td>
<td>0.356</td>
<td></td>
<td>5.1</td>
<td>0.000*</td>
</tr>
<tr>
<td>Sleep latency/ min</td>
<td>7.5</td>
<td>0.000</td>
<td>Self-reported:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N3 latency</td>
<td>3.7</td>
<td>0.003</td>
<td>Quality of rec. sleep</td>
<td>7.0</td>
<td>0.000*</td>
</tr>
<tr>
<td>REM latency</td>
<td>3.0</td>
<td>0.012</td>
<td>PSG affected sleep</td>
<td>0.9</td>
<td>0.465</td>
</tr>
<tr>
<td>N1</td>
<td>2.7</td>
<td>0.019</td>
<td>Habitual/usual sleep duration</td>
<td>1.3</td>
<td>0.239</td>
</tr>
<tr>
<td>N1%</td>
<td>4.2</td>
<td>0.001</td>
<td>Sleep duration different from usual</td>
<td>26.9</td>
<td>0.000</td>
</tr>
<tr>
<td>N2</td>
<td>65.9</td>
<td>0.000</td>
<td>Sleep quality different from usual</td>
<td>4.0</td>
<td>0.002*</td>
</tr>
<tr>
<td>N2%</td>
<td>0.2</td>
<td>0.756</td>
<td>Habitual sleep medication</td>
<td>0.7</td>
<td>0.613</td>
</tr>
<tr>
<td>N3</td>
<td>4.7</td>
<td>0.000</td>
<td>Habitual sleep quality</td>
<td>1.4</td>
<td>0.224</td>
</tr>
<tr>
<td>N3%</td>
<td>0.4</td>
<td>0.835</td>
<td>HAD anxiety</td>
<td>3.0</td>
<td>0.011</td>
</tr>
<tr>
<td>REM/10 min</td>
<td>51.7</td>
<td>0.000</td>
<td>HAD depression</td>
<td>0.9</td>
<td>0.487</td>
</tr>
<tr>
<td>REM%</td>
<td>10.3</td>
<td>0.000</td>
<td>Lights out</td>
<td>14.2</td>
<td>0.000</td>
</tr>
<tr>
<td>Fast spindles/hr N2</td>
<td>1.2</td>
<td>0.296</td>
<td>Time of awakening</td>
<td>8.8</td>
<td>0.000</td>
</tr>
<tr>
<td>Slow spindles/hr N2</td>
<td>1.2</td>
<td>0.303</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. AHI, apnea-hypopnea index; PSG, polysomnography; HAD, Hospital Anxiety and Depression Scale; WTSP, wake time during total sleep period; REM, rapid eye movement sleep.
All significant F-ratios showed a significant linear contrast at <0.002, except for HAD depression (with p = 0.011).
*F-ratios with a significant quadratic contrast at p < 0.05.

increasing TST, whereas the number of awakenings/hr, the number of arousals/hr, N1%, sleep latency and REM intensity decreased. In addition, long sleep was more associated with sleep duration being perceived as a “usual” sleep duration and with better quality. Long sleep was associated with earlier time for “lights out” and later time of awakening.

There are no previous studies of the cross-sectional relation between TST and polysomnographical variables to make comparisons with. However, in a previous study on the same study population we found that poor subjective sleep quality was related to short TST, low sleep efficiency, more awakenings and fewer arousals (Åkerstedt, Schwarz, Gruber, Theorell-Haglow, & Lindberg, 2018), which are in line with the present results, even if TST was never related to other PSG variables. Similarly, a meta-analysis has demonstrated impaired sleep continuity, and reduced SWS% and REM%, as well as lower TST (Baglioni et al., 2014), in insomniacs. The present study, however, directly links short sleep to (dis)continuity variables (sleep efficiency, sleep latency, awakenings, arousals and short segments of continuous sleep), but also to increased N1%, which may be interpreted as an indicator of poor sleep. The observation of N1 in minutes being lower for short sleep probably reflects that shorter TST gives less room for N1 or any other sleep stage.

One could also compare the present findings with what is seen in experimental partial sleep deprivation; that is, increased sleep continuity and N3%, and decreased REM% and N1% across 5 days of only 4 hr sleep per day (Åkerstedt, Kecklund, Ingre, Lekander, & Axelsson, 2009). The 4-hr duration corresponds to the shortest sleep in the present study. A similar reduction of N1 and REM was seen in another study of partial sleep deprivation (Brunner, Dijk, Tobler, & Borbély, 1990). In addition, a sudden delay of bedtime by 4 hr led to decreased N3 and REM in minutes and in percent REM, but N3% increased significantly (Åkerstedt et al., 2018). Although TST in that study (3 hr) was similar to that of the short sleepers in the present study, the delay of bedtime was only about 45 minutes; that is, very modest partial sleep deprivation.

The present study is, however, quite different from studies of experimental sleep deprivation and the results suggest a temporary impairment of sleep in the short sleepers on the recorded night. This interpretation of the main results seems relatively straightforward, in the sense that not only sleep continuity and N1% were affected, but also the morning sleep quality ratings showed considerable impairment. In contrast, habitual sleep quality bore no relation to PSG, indicating that the recorded sleep may not have been representative of habitual sleep quality. Thus, the short sleepers do not seem to have been insomniacs, but rather individuals who experienced a transient bout of poor sleep. This is supported by the ratings of the recorded sleep having been shorter and poorer “than usual”.

Apparently, the recorded sleep deviated from normal sleep quality and sleep duration for unknown reasons. One reason could have been that the sleep recording itself had disturbed sleep, but this was not reported by the participants. Neither was TST associated with ratings of “anything special having happened during the day that
may have affected sleep”. Still, there may have been effects of the recording or of other external factors without the participants recognizing this. At present, this question remains unanswered and should be a topic for further research. Furthermore, the lack of association with depression and the weak association with anxiety suggests that the TST groups did not differ emotionally. Taken together, the results suggest that a short sleep may be an indicator of a temporary poor sleep, at least in individuals who are otherwise normal sleepers and who have their sleep recorded in a normal daily life setting outside the sleep laboratory.

The result also suggests that the TST of a recorded night of sleep may not be representative of reports of “habitual” sleep duration or quality, at least not when recorded under real-life circumstances in a representative sample of women. Previous studies of long and short sleep have carried out recordings under well-controlled laboratory conditions without any obvious curtailment of sleep (and in young male students) (Aeschbach et al., 1996; Benoit et al., 1983; Webb & Friel, 1970). This is unlikely to produce results representative of a real-life sleep episode. This raises the question of what habitual sleep duration represents. One possibility is that it represents an integration of “normal” sleep across long periods of time and that occasional impaired sleep episodes may not be included in this estimate.

Apart from sleep continuity, REM sleep in minutes strongly increased with sleep duration. This is a finding that goes back to the time of the first polysomnographical studies (Aeschbach et al., 1996; Benoit et al., 1983; Webb & Friel, 1970). This probably reflects the increase in REM presence with increasing sleep duration, but may also reflect a sensitivity of REM to poor sleep continuity as REM% is reduced in insomniacs (Baglioni et al., 2014). Furthermore, recent findings have identified “restless REM” (fragmented REM with a high frequency of eye movements) as an important marker of insomnia (Wassing et al., 2016). Interestingly, in the present study REM intensity (the amplitude of rapid eye movements) decreased markedly with increased TST. The interpretation of this finding is not clear and needs replication. The finding of a higher REM density in the study by Wassing et al. (2016), which together with REM arousals was strongly associated with slow dissolution of emotional distress, concerned insomniacs and it is not clear if this was related to TST or other PSG parameters.

The association between N3 duration and TST was expected, because longer sleep would allow more room for N3. However, percent N3 was not related to TST, presumably because N3 is present only in the first cycles of sleep (Williams, Karacan, & Hursch, 1974), and with each minute of sleep beyond that the denominator of the sleep percentage ratio would increase, thus decreasing the percentage. Our new variable, delta dominance, was particularly low in the group with shortest TST. This variable was developed in an attempt to standardize for differences between individuals in anatomical structures that might affect the EEG amplitude in the 0.5–4-Hz band. The impression is, therefore, that slow wave sleep was particularly impaired in the group with the shortest sleep (and with the highest level of sleep discontinuity), despite the observation of no association between TST and N3%.

This is a study with a large sample and with good representativeness of women in a Swedish city. However, there are limitations. There was only one night being recorded and it would have been of interest to study several nights in order to investigate whether short/poor sleep was a temporary phenomenon. Only women were studied and it would have been an advantage to include also men. The present study also failed to find an explanation for impaired sleep for the short sleepers. This seems to be an important topic for future studies. Presently, we assume that short/poor sleep was temporary, based on the report of such sleep being unusually poor and short, and on habitual sleep quality or duration having no association with the PSG measures.

In conclusion, the present study demonstrates that short TST in a recorded sleep in a real-life context may be an indicator of poor objective (and subjective) sleep quality for that particular sleep episode. Because individuals clearly perceived this reduction, it appears that self-reports of poor sleep quality often may be seen as indicators of poor sleep quality.

CONFLICT OF INTEREST

No author has reported a conflict of interest.

AUTHOR CONTRIBUTION

Eva Lindberg and Jenny Theorell-Haglöw initiated the study, led the data collection and contributed to the writing of the paper. Torbjörn Åkerstedt wrote the paper and analyzed the data. Johanna Schwarz coordinated the data management and contributed to the writing of the paper. Georg Gruber led the computer-based sleep scoring and contributed to the writing of the paper.

ORCID

Torbjörn Åkerstedt https://orcid.org/0000-0001-8049-8504
Johanna Schwarz https://orcid.org/0000-0001-9873-2506
Jenny Theorell-Haglöw https://orcid.org/0000-0003-2819-2799

REFERENCES


