

<https://helda.helsinki.fi>

---

The adapted American Academy of Sleep Medicine sleep scoring criteria in one month old infants : A means to improve comparability?

Satoomaa, Anna-Liisa

2016-02

---

Satoomaa , A-L , Saarenpaa-Heikkila , O , Paavonen , E J & Himanen , S-L 2016 , ' The adapted American Academy of Sleep Medicine sleep scoring criteria in one month old infants : A means to improve comparability? ' , Clinical Neurophysiology , vol. 127 , no. 2 , pp. 1410-1418 . <https://doi.org/10.1016/j.clinph.2015.08.013>

---

<http://hdl.handle.net/10138/223868>

<https://doi.org/10.1016/j.clinph.2015.08.013>

---

publishedVersion

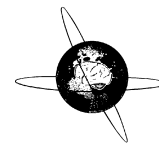
---

*Downloaded from Helda, University of Helsinki institutional repository.*

*This is an electronic reprint of the original article.*

*This reprint may differ from the original in pagination and typographic detail.*

*Please cite the original version.*



## The adapted American Academy of Sleep Medicine sleep scoring criteria in one month old infants: A means to improve comparability?



Anna-Liisa Satomaa<sup>a,\*</sup>, Outi Saarenpää-Heikkilä<sup>b</sup>, E. Juulia Paavonen<sup>c,d</sup>, Sari-Leena Himanen<sup>a,e</sup>

<sup>a</sup> Department of Clinical Neurophysiology, Tampere University Hospital, Medical Imaging Centre and Hospital Pharmacy, Pirkanmaa Hospital District, Tampere, Finland

<sup>b</sup> Tampere University Hospital, Department of Pediatrics, Tampere, Finland

<sup>c</sup> National Institute for Health and Welfare (Child and Adolescent Mental Health), Helsinki, Finland

<sup>d</sup> Children's Hospital, Child Psychiatry, University of Helsinki and Helsinki University Central Hospital, Finland

<sup>e</sup> School of Medicine, University of Tampere, Tampere, Finland

### ARTICLE INFO

#### Article history:

Accepted 18 August 2015

Available online 2 September 2015

#### Keywords:

Infant polysomnography

Sleep stage scoring

Infant sleep

### HIGHLIGHTS

- 84 one-month-old infants' polysomnographies (PSGs) were scored by the adapted AASM sleep scoring rules.
- The rules were feasible, but the sleep parameters showed wide individual variation.
- Comparison of data to literature underlines the need for uniform scoring methods.

### ABSTRACT

**Objective:** The lack of standards induces variability in the sleep staging of infants less than two months of age. We evaluated the feasibility of the 2012 AASM sleep scoring rules for healthy one month old infants.

**Methods:** 84 polysomnographies were scored into sleep stages with the adapted AASM criteria. The acquired sleep parameters were compared with the parameters in the literature. In addition the effect of age on sleep was studied.

**Results:** The two independent scorers achieved substantial agreement by using the adapted AASM criteria. The infants' sleep parameters showed marked variability. The amount of active sleep was 36.7% (mean, range 21.3–54.1%), quiet sleep 41.5% (30.3–57.7%) and indeterminate sleep 21.6% (9.7–36.0%). With age sleep became more continuous, but the sleep stage percentages did not change. Our sleep parameters differed clearly from the parameters presented in the literature.

**Conclusions:** The adapted scoring rules were reproducible. This encourages their use in clinical practice, as no uniform recommendations exist.

**Significance:** Normal values are essential in pediatric sleep medicine and the individual variability in the sleep parameters of healthy infants advocates the standardisation of scoring methods. Here we present sleep stage normative values for one month old infants based on the AASM scoring criteria.

© 2015 International Federation of Clinical Neurophysiology. Published by Elsevier Ireland Ltd. All rights reserved.

## 1. Introduction

Newborns and young infants spend most of their time sleeping. Cumulating evidence suggests that the quality of early sleep can affect later health and development. For example, more mature

\* Corresponding author at: Department of Clinical Neurophysiology, Tampere University Hospital, Medical Imaging Centre and Hospital Pharmacy, Pirkanmaa Hospital District, P.O. Box 2000, 33521 Tampere, Finland. Tel.: +358 3 31165095; fax: +358 3 31164199.

E-mail address: [anna-liisa.satomaa@pshp.fi](mailto:anna-liisa.satomaa@pshp.fi) (A.-L. Satomaa).

sleep patterns in infancy relate to better level of mental development at 6 months of age (Gertner et al., 2002). On the other hand, sleep disordered breathing, as a common sleep disrupting condition, is known to cause negative behavioral and cognitive consequences (O'Brien and Gozal, 2002; Halbower et al., 2006). The duration of sleep relates to infants' physical growth already during the first months of life (Tikotzky et al., 2010).

The normal development of an infant and the age specific changes in the infant's sleep are parallel phenomena. During the neonatal period these developmental changes are most profound and they include the rapid maturation of sleep electroencephalography

(EEG), which allows a better distinction of the sleep stages. In the first months of life, sleep EEG does not yet differentiate into mature sleep stages. Instead, NREM sleep and REM sleep are termed quiet sleep (QS) and active sleep (AS), respectively. Young infants also present decreasing amounts of indeterminate sleep (IS) which has features of both AS and QS. The sleep EEG of QS changes the most during the first two months: the neonatal tracé alternant pattern would be gradually replaced by continuous high voltage slow activity, and sleep spindles and more mature slow wave sleep would appear. This development allows the QS to be divided in more mature NREM stages by the age of three months (Dreyfus-Brisac, 1970; Coons and Guilleminault, 1982; Louis et al., 1992).

The maturation of the sleep EEG is paralleled by the development of more organized and longer sleep cycles. The relative proportions of sleep stages change; AS and IS decrease, QS increases. The emergence of circadian and homeostatic regulation consolidates sleep and changes the temporal distribution of the sleep stages within a sleep period and according to the time of day (Dittrichová, 1966; Coons and Guilleminault, 1982; Hoppenbrouwers et al., 1982; Schechtman et al., 1994; Hellström-Westas et al., 2003; So et al., 2007). At the same time, the mechanisms of arousal (McNamara et al., 1996; Montemitto et al., 2008) and the regulation of other physiological sleep related phenomena would evolve.

As the sleep EEG and the structure of sleep are dependent on the conceptional age, the normality of sleep must always be evaluated as a function of the infant's age (Parmelee and Stern, 1972). Because the age-specific reference values are essential for the interpretation of pediatric polysomnographies (PSGs), to obtain the values for each age group, numerous studies have been carried out over the past decades. In the past, the methods to define the age-specific reference values have varied from behavioral observation (Aserinsky and Kleitman, 1955; Dement and Kleitman, 1957; Anders and Sostek, 1976; Anders and Keener, 1985) to actigraphy based studies (So et al., 2007) and to cardiorespiratory parameter based scoring (Haddad et al., 1987). Although behavioral correlates continue to be helpful in the scoring of sleep in young infants, the sleep staging has begun to rely more on the EEG. However, since there have been no universally accepted sleep scoring criteria for infants, many researchers have had to create or modify their own EEG-based scoring criteria with more or less emphasis on behavioral and observational aspects (Coons and Guilleminault, 1982; Hoppenbrouwers et al., 1988; Kirjavainen et al., 2004; Hoppenbrouwers et al., 2005; Montemitto et al., 2008). Because of the different methods, it is difficult to adapt the results of the previous normative studies to the clinical work.

The current AASM scoring manual (Berry et al., 2012) is intended for infants from two months of age on. For infants younger than two months, the AASM Pediatric Task Force (Grigg-Damberger et al., 2007) recommends the Anders manual (Anders et al., 1971), in which the observational data are underlined in addition to polysomnographic parameters. The AASM 2012 criteria, by contrast, do not require visual observation, and the criteria are in routine use for older infants and adults at our laboratory. The main aim of this study was to obtain sleep parameter values for normal one month old infants by applying the AASM scoring manual as closely as possible. The resulting sleep parameters were compared to the previously presented normal values in literature.

## 2. Methods

This study is part of a large birth cohort (Child Sleep) which consists of 1671 infants born at Tampere University Hospital (TAYS) during 4/2011–2/2013. The Child Sleep is a multidisciplinary research project that aims to study various aspects of sleep

in early childhood as well as the effect of sleep on the children's later well-being and development. The research utilizes questionnaires as well as the genetic data. The present study consists of a subgroup of this cohort, in which the participating infants underwent an ambulatory over-night PSG in three different ages (1, 8, and 24 months of age).

The Child Sleep families were recruited prenatally at their local maternity clinics to take part of the study. Systematically all the families, who were eligible for the PSG sub study according to pre defined inclusion and exclusion criteria, were personally asked to participate in the PSG sub study postnatally on their stay at the maternity ward. The inclusion criteria for the PSG study were: healthy, full-term and uneventful birth (conceptional age 38 weeks or more), Apgar score  $\geq 8$  at one minute and birth weight  $\geq 2500$  g. The mothers were healthy, they could not have had been on any medication affecting central nervous system during their pregnancy, and they would have attended to their routine controls at the local maternity clinic. Originally, 92 families who fulfilled the inclusion criteria agreed to participate in the PSG study. Closer to the recording, four families withdrew. The remaining 88 singleton infants were born at the conceptional age of 38–42 weeks. The first PSGs, at the age of one month (range 3.1–7.9 weeks), were recorded between May 2011 and January 2013. All the recruited parents gave their written informed consent.

On the scheduled measurement night, a PSG technician and a medical physicist for technical support arrived in the families' homes approximately two hours before the family's usual bed time. The same PSG technician and two different medical physicists acquired all recordings. After the electrodes and sensors were placed and the function of all channels was checked on a laptop, the recording was started. The parents were asked to keep notes of the events (e.g. feeding and diaper changing) during the night. The following morning the PSG technician returned to the family and removed the electrodes. The quality and specific events of the measurement night were briefly discussed and the parents were asked to evaluate whether the measurement night had been a typical night. 33% of the families felt that the PSG night had been worse than usual, and 23% of the families found the night better than a typical night. The majority (45%) of families thought that the night had been typical in quality.

### 2.1. Participants

In total, 88 one-month-old infants were recorded. Three recordings (3%) were lost because of a technical failure. One infant was excluded from the statistical analyses because, inadvertently, she was younger than 42 weeks of conceptional age ( $41 + 2$ ) at the time of the recording. The quality of the remaining 84 recordings was adequate for sleep stage scoring.

In terms of respiratory analysis, there were some technical problems and nine recordings were excluded from analyses concerning oxygen saturation. In the excluded nine recordings the apneas were scored on the bases of the event duration (duration  $\geq 2$  breaths during baseline breathing) even though there was no associated cortical arousal. Undoubtedly, this slightly overestimates the apnea indices in these nine infants. Nevertheless, their apnea-hypopnea indices (AHIs) and obstructive AHIs (OAHIs) were not statistically different from those infants with functioning oximeters ( $p = 0.17$  and  $p = 0.43$ , respectively). The thermistor failed in four recordings, which only inhibited scoring obstructive events.

After the exclusions, the study sample consisted of 84 infants (42 boys, 50%). The average of the gestational age at birth was 40.2 weeks (38.3–42.0 weeks). At the time of the recording the average of the conceptional age was 44.7 weeks (42.3–48.4 weeks). The study sample for the pulse oximeter-derived parameters

consisted of 75 infants, whose mean conceptional age was 44.6 weeks (42.3–48.4 weeks) at the recording.

## 2.2. Recordings and visual analysis

The ambulatory PSGs were obtained using the Embla Titanium system. The following signals were recorded: 6 channels of EEG (F4-A1, C4-A1, O2-A1, F3-A2, C3-A2, O1-A2), electro-oculography (EOG) (right and left), submental electromyography (EMG), oxygen saturation (pulse oximeter, Nonin, with two-beat averaging time) with waveform, thoracoabdominal inductance plethysmography, diaphragmatic and abdominal EMG, Emfit mattress sensor and electrocardiography (ECG). Airflow was measured by oronasal thermistor (Dymedix). The sampling frequencies for the EEG, EOG, EMG and ECG were 256 Hz; for the Emfit mattress 128 Hz, for the thermistor 64 Hz, for the plethysmographies 32 Hz, and for the pulse oximetry 16 Hz. The nasal pressure transducer was omitted from the protocol in order to minimize the sleep disturbing effect of the recording equipment (Goodwin et al., 2001).

EEG was recorded with Ag–AgCl cup EEG electrodes, which were attached with water soluble paste, their cables were braided together, and the electrodes were covered by a mesh cap. The EOGs and chin EMG were recorded with single use self adhesive electrodes. After all the electrodes and sensors were placed, their function was checked on a laptop, electrode impedances <5 kOhm were accepted. The filter settings for EEG were 0.5–70 Hz. The EEG was considered acceptable in quality if at least three of the six channels remained readable throughout the recording. In six recordings three EEG traces were lost (usually due to the detachment of A1 or A2), in seven recordings one EEG electrode became loose during the night, and in one recording two traces were lost. Hence, none of the PSGs had to be excluded because of EEG quality alone. The mean amplitude levels during QS were ad 142 uV and during AS ad 46 uV.

The PSGs were scored into the sleep stages in 30-s epochs with Somnologica Studio 5.0 software by two independent, experienced clinical neurophysiologists (ALS, SLH). The two scorers had an inter-scorer agreement of 80.6% and kappa score of 0.73 indicating substantial agreement (Landis and Koch, 1977). The agreements and differences in scored epochs are presented in Table 1. The agreement by sleep stage was >80% in AS, QS and wakefulness (W), and 46% in IS. The most common mismatch between the sleep stages was IS-AS, which comprised 27.4% of all mismatches. The other mismatches involving IS (IS-AS and IS-W) constituted

14.4% and 14.3% of all mismatches, respectively, while the other mismatches (W-AS, AS-QS and W-QS) accounted for only 10%, 5%, and 2% of all mismatches, respectively. Afterwards, any major scoring differences were discussed and a consensus scoring was established for statistical analyses. Table 1 also presents the agreements and differences of the two scorers compared to the final consensus scoring. The total agreement of scorer 1 and the consensus scoring was 82.9%, kappa score 0.76. The total agreement for scorer 2 was 79.7% and kappa score 0.72. For both scorers, the agreement with the consensus scoring was clearly lowest in IS, whereas in the other sleep stages and wakefulness the agreement with the consensus scoring was better.

The sleep stages were scored according to the AASM 2012 scoring manual (Berry et al., 2012) as closely as possible, but due to the infants' age-related sleep-EEG features, some modifications were implemented. The infant's state was scored as active sleep (AS), quiet sleep (QS), indeterminate sleep (IS) or wakefulness (W). AS, W and QS were scored according to the REM sleep, wakefulness and deep sleep rules (R, W and N3 rules) in the AASM manual, respectively.

QS was scored when the EEG presented either 1 a high voltage slow (HVS) pattern (continuous, moderately rhythmic, high voltage,  $\geq 75$  uV, 0.5–2 Hz activity) in any EEG channel for at least 20% of an epoch's duration or 2 a tracé alternant pattern. Because the slow wave activity (SWA) of N3 sleep first appears from two months of age on, and because the slow activity in young infants does not necessarily represent the same generating mechanisms as in an older child or an adult (Schechtman et al., 1994; Jenni et al., 2004), the slow activity was termed HVS instead of SWA, and the HVS on any EEG channel was accepted. In infants, the amplitude of HVS is generally high, >100 uV.

AS was scored when the EEG did not present HVS or tracé alternant (usually mixed frequency activity of low to moderate amplitude) and the chin EMG was low except for short transient bursts. In stage AS, the eye movements were rapid (rise time <0.5 s), but they were not required to continue the stage.

Stage W was scored when rapid eye movements were associated with sustained EMG tone with activity bursts or gross movements, or when movement artefact covered at least 50% of the duration of an epoch followed or preceded by a wake epoch. IS was scored when an epoch could not be scored as AS, QS or W. These included typically the state transition segments.

In the AASM scoring rules, no other PSG parameters in addition to EEG, chin EMG, and EOG, are recommended to be taken into account. However, even in the adult population, the chin EMG in N3 sleep is of variable amplitude and sometimes as low as in stage R sleep (Berry et al., 2012). In young infants, there is a known lack of concordance between the EMG tone, respiratory pattern and other PSG signals (Dreyfus-Brisac, 1970; Parmelee et al., 1972; Schloon et al., 1976). Moreover, in young infants, the EEG presents ample slow EEG activity irrespective of sleep stage (Pereda et al., 2006; Sankupellay et al., 2011). That is why the respiratory pattern or the variability of heart rate were used to support the decision making. Irregular breathing pattern and variable heart rate were considered to favor AS. Regular breathing with little heart rate variability and sustained moderate EMG tone were considered to support QS. Breathing pattern was considered regular, if there were no visible accelerations/decelerations typical for AS (R) during continuous breathing. Regular heart rate was visually recognized when the ECG lacked the typical accelerations of AS.

The duration of each sleep state divided by total sleep time (TST) were expressed as AS%, QS%, IS%.

The cortical arousals were visually scored according to the current guidelines (The IPWG, 2005; Grigg-Damberger et al., 2007; Berry et al., 2012). Hence, cortical arousals were scored when the EEG background frequency presented an abrupt change of at least

**Table 1**  
The differences and agreements in sleep staging between the two independent scorers and the final consensus scoring.

	W	AS	IS	QS
Scorer 1 vs. consensus				
W	88	4.8	6.3	0.4
AS	7.8	75.7	15.6	0.3
IS	11.8	18.8	63.4	6.8
QS	0.6	0.8	7.4	91
Scorer 2 vs. consensus				
W	81.1	8.6	9.7	0.6
AS	2.3	81.9	13.9	1.8
IS	11	27.3	49.6	12.4
QS	0.9	2.4	6.3	90.1
Scorer 1 vs. scorer 2				
W	82.4	7.7	8.9	1
AS	2.2	80.1	14.7	3
IS	11.2	26.5	46.2	16.1
QS	0.8	2.4	6.8	89.4

All numbers are in percents. W = wakefulness; AS = active sleep; IS = indeterminate sleep; QS = quiet sleep.

1 Hz for a minimum of 3 s. A decremental EEG-response was accepted as an arousal, a frequency shift in delta range was not. To accompany this EEG change, two of the following changes were required for scoring arousals in AS: (1) a gross body movement artefact, (2) increase in the heart rate (at least 10% from the baseline values) or (3) an increase in the submental EMG amplitude. In QS and IS, two of the following criteria were required to score arousals in addition to the EEG change: (1) a gross body movement, (2) change in the heart rate or (3) a change in the breathing pattern (frequency and/or amplitude) including a single sigh. An arousal with a duration of  $\geq 30$  s, or an arousal followed by a wake epoch was equaled as an awakening. The number of awakenings divided by TST and the number of awakenings together with cortical arousals divided by TST were expressed as awakening index and arousal index, respectively.

The respiratory events were scored according to the pediatric rules in the AASM scoring manual (Berry et al., 2012). The number of all apneas and hypopneas divided by TST equaled apnea–hypopnea index (AHI). The number of all obstructive events (obstructive and mixed apneas and obstructive hypopneas) divided by TST was expressed as obstructive apnea–hypopnea index (OAHI). The number of central apneas and central hypopneas divided by TST were expressed as central apnea index (CAI) and hypopnea index, respectively. More detailed respiratory parameters and discussion are outside the scope of this paper.

### 2.3. Effect of age in the sleep parameters

To study the effect of age on the sleep parameters, the infants were divided in two age groups: the infants less than 44 weeks of age at the recording and the infants 44 weeks or older. Thirty-one (37%) of the 84 infants were less than 44 weeks of conceptional age at the time of the recording (mean age 43.2 weeks, range 42.3–43.9) leaving 53 of the infants to the older age group (mean 45.4, range 44.0–48.4). In the younger group there were 14 (45%) boys; in the older group 37 (49%) were boys. The division to age groups also enabled comparisons with previous studies.

### 2.4. Comparison to the previous studies

The division to age groups also enabled comparisons with previous studies evaluating sleep parameters of small infants. We compared the sleep parameters in our data to six previously published studies (Coons and Guilleminault, 1982; Anders and Keener, 1985; Hoppenbrouwers et al., 1988, 2005; Kirjavainen et al., 2004; Montemitto et al., 2008). The compared studies were selected as they covered the same age group of full-term and healthy infants and presented the descriptives of the sleep parameters (means and standard deviations) numerically. The number of the participants and their ages, the recorded signals and the applied scoring criteria are summarized in Table 2.

Among the six compared studies, none applied the same exact scoring criteria. The studies by Anders and Keener (1985), Coons and Guilleminault (1982) and Montemitto et al. (2008) applied the Anders manual (Anders et al., 1971), but all had made some modifications to suit their set-up. The Anders manual divides the infant's state into quiet (NREM) sleep, AS, wakefulness, and IS. The scoring of these states requires information on movements, eye closure, eye movements, EEG pattern, chin EMG tone, and the regularity of breathing. The study by Montemitto et al. (2008) followed these rules to most detail with continuous observation during the study night. Coons and Guilleminault (1982) had modified the rules somewhat, as they did not have continuous observation or EOG. Anders and Keener (1985) made use of the behavioral criteria of the Anders manual (movements, facial

expressions, vocalization, eye movements, eye closure, and, if visible on the video, the regularity of breathing).

The study of Hoppenbrouwers et al. (1988) applied the Hoppenbrouwers criteria (Hoppenbrouwers, 1987), which divide the infant state into AS, QS, W and IN based on chin EMG, EOG, respiratory pattern, movements or twitches, eye closure, vocalization, and EEG. The minimum number of these parameters required to score a state are defined: AS requires four criteria, QS and W require three. IN is scored if the criteria for QS, AS or W are not fulfilled.

The Guilleminault and Souquet criteria (Guilleminault and Souquet, 1982) applied in the study by Kirjavainen et al. (2004) are meant for infants older than three months of age. According to the criteria an infant's state is scored as W, sleep onset, stages 1–2, stages 3–4, and REM. The scoring makes use of EEG, EOG, chin EMG and behavioral criteria. Because the infants in the study by Kirjavainen et al. (2004) were younger than three months of age and because observational data was not available, the criteria were modified. The stages were named REM, deep NREM, light NREM and W, and their scoring relied mainly on EEG criteria, respiratory pattern, chin-EMG tone; movements were interpreted from movement artifacts.

Hoppenbrouwers et al. (2005) applied the CHIME scoring rules (Crowell et al., 1997). The rules are based on EEG, respiratory variability, eye movements, and body movements and they are detailed in terms of how many criteria must be fulfilled to score a stage. The state is scored as AS or QS. If the majority of AS or QS parameters are not fulfilled, the state is scored as indeterminate, while the state W is scored based on movements and annotations.

### 2.5. Statistics

The normality of the data was tested with Kolmogorov–Smirnov and Shapiro–Wilk tests, which indicated that all the tested sleep parameters, except for QS%, were normally distributed. The Mann–Whitney *U*-test was used in the between-group comparisons (for example age group differences). The correlations were tested with Spearman's rho test. These statistical comparisons were performed with the SPSS software (v 21). To compare the obtained sleep parameters with the parameters presented in the literature, we applied the one sample *t*-test, which takes into account the presented means, standard deviations, and sample sizes (GraphPad Software, Inc., 2015). *p*-values <0.05 were considered statistically significant in statistical analyses.

## 3. Results

### 3.1. All infants

The sleep quality parameters and the basic respiratory indices are presented in Table 3. During nocturnal sleep the 84 infants presented in average 36.7% AS, 41.5% QS and 21.6% IS (Table 1). The awakening index was 8.5/h and arousal index 19.3/h. There was considerable individual variation in the sleep state percentages and especially in the arousal indices. Based on the  $\pm 2$  SD values, the reference limits for %AS, %QS, %IS, awakening index and arousal index in this population of infants were: 21.3–54.1%, 30.3–57.7%, 10.0–33.2%, 4.4–12.6/h and 11.9–27.0/h, respectively.

The correlation of arousal index and the awakening index with the sleep stage percentages was tested. The only statistically significant correlation was found between the percentage of IS and the awakening index ( $r = 0.245$ ,  $p = 0.025$ ); the infants with a higher awakening index presented statistically more IS, but the correlation was very weak.

**Table 2**  
Compared studies.

Study	n	Age at recording	In-laboratory/ at home	Over-night/ 24 h	Recorded signals	Scoring criteria
Coons and Guillemainault (1982)	10	3 weeks	In-lab	24 h	EEG <sup>a</sup> (number of electrodes not specified), chin-EMG <sup>a</sup> , respiratory <sup>a</sup>	Anders et al. (1971) (EEG, behavioral)
Anders and Keener (1985)	40	4 weeks	Home	Over-night	Time-lapse video	Anders et al. (1971) (observational)
Hoppenbrouwers et al. (1988)	10	1 month	In-lab	Over-night	EEG (two channels), chin-EMG, EOG <sup>a</sup> , respiratory	Hoppenbrouwers (1987) (EEG, behavioral; observation supplementary)
Kirjavainen et al. (2004)	23	1 month	Home	Over-night	EEG (C3, C4), chin-EMG, EOG, respiratory	Hoppenbrouwers (1987)
Hoppenbrouwers et al. (2005)	60	44–49 weeks conceptional age	In-lab	Over-night	EEG (C3, C4), chin-EMG, EOG, respiratory	Guilleminault and Souquet (1982) (EEG, behavioral)
Montemitro et al. (2008)	10	3 weeks	In-lab	Over-night	EEG (1 central and 1 occipital lead), chin-EMG, EOG, respiratory	Crowell et al. (1997) (EEG, behavioral, observational annotations)
						Anders et al. (1971) and The IPWG (2005) (EEG, behavioral, continuous observation)

<sup>a</sup> EEG = electroencephalography; EMG = electromyography; respiratory = at least one sensor to detect respiratory movements; EOG = electro-oculography.

**Table 3**  
The sleep quality and the respiratory parameters of 84 infants. Pulse oximetry data in 75 infants.

	Mean	Range	SD	±2 SD
Analyzed period (TIB) (min)	718	540–915	81.9	
TST (min)	476.9	305–628	58.5	
AS%	36.7	21.3–54.1	6.9	22.8–50.5
QS%	41.5	30.3–57.7	5.2	31.2–51.9
IS%	21.6	9.7–36.0	5.8	10.0–33.2
Awakening index (n/h)	8.5	4.5–13.3	2.1	4.4–12.6
Arousal index (n/h)	19.3	12.1–29.0	3.7	11.9–27.0
AHI (n/h)	13.7	0.5–81.7	14.1	0–41.7
OAH1 (n/h)	0.0	0–0.3	0.0	0–0.1
ODI3 (n/h)	20.9	0.0–79.5	16.9	0–54.7

ODI3 = Oxygen desaturation index, the number of oxygen desaturations  $\geq 3\% \times 60$ /TST.

In the morning following the recording night, the parents evaluated the quality of sleep compared to a typical night. We tested whether the sleep parameters of the 28 (33%) infants who were reported to have slept better than usual differed from the 19 (23%) infants who were reported to have slept worse than usual. The sleep stage percentages or arousal and awakening indices did not statistically differ between these groups of parent-reported sleep quality (all  $p$ -values  $\geq 0.05$ ).

### 3.2. Age group analyses

In order to study the effect of age on sleep quality, the 84 infants were divided in two groups according to conceptional age. The per-

centages of the sleep stages were not statistically different between the age groups  $<44$  weeks and  $\geq 44$  weeks (Table 4). However, the younger infants' sleep was more discontinuous – both the number of awakenings/h (9.4 vs. 7.9, respectively) and arousal index (20.7 vs. 18.4) were significantly ( $p < 0.05$ ) higher in the younger age group. The younger infants had significantly higher total AHI and OAH1 ( $p = 0.03$  and  $p = 0.04$ , respectively), although the incidence of obstructive apneas was very low in both age groups.

### 3.3. Comparisons to previous normal values

The sleep parameters of the data were compared to six previously published studies (Coons and Guillemainault, 1982; Anders and Keener, 1985; Hoppenbrouwers et al., 1988; Kirjavainen et al., 2004; Hoppenbrouwers et al., 2005; Montemitro et al., 2008). Nearly all comparisons yielded significant differences in the amounts of sleep stages between our data and the previous studies; the only non-significant differences were found in the IS % in the study by Hoppenbrouwers et al. (1988), QS% in the study by Coons and Guillemainault (1982), and in the combined IS + AS proportion in the study by Montemitro et al. (2008). Except for the study of Coons and Guillemainault (1982), in our data, the QS % was significantly higher than in the compared studies. The proportions of the other stages (AS and IS), however, were either significantly higher or lower: we found more AS than Coons and Guillemainault (1982), Hoppenbrouwers et al. (2005) and Kirjavainen et al. (2004). Our percentage of IS was higher than in Hoppenbrouwers et al. (2005) (group 1) and Hoppenbrouwers et al. (2005). Table 5 presents the comparisons.

**Table 4**  
The sleep quality and respiratory parameters in the age groups  $<44$  weeks ( $n = 31$ ) and  $\geq 44$  weeks ( $n = 53$ ) of conceptional age at the time of the recording. Pulse oximetry (ODI3) data in 27 infants aged  $<44$  weeks and 48 infants aged  $\geq 44$  weeks. Statistically significant  $p$ -values are in bolded italics.

	$<44$ weeks ( $n = 31$ )				$\geq 44$ weeks ( $n = 53$ )				$p$
	Mean	Range	SD	±2 SD	Mean	Range	SD	±2 SD	
Age (weeks)	43.2	42.3–43.9	0.48		45.4	44.0–48.4	0.87		
AS%	36.7	26.5–47.5	6.2	24.2–49.2	36.7	21.3–54.1	7.3	22.2–51.3	0.948
QS%	41.2	34.9–54.6	4.1	32.9–49.3	41.8	30.3–57.7	5.7	30.4–53.2	0.423
IS%	22.5	9.7–36.0	6.1	10.2–34.7	21.1	12–35.4	5.5	10.1–32.2	0.207
Awakening index	9.4	5.9–13.3	2	5.5–13.4	7.9	4.5–12.3	1.9	4.2–11.7	<b>0.001</b>
Arousal index	20.7	13.5–29.0	3.7	13.4–28.0	18.4	12.1–26.0	3.4	11.6–25.2	<b>0.004</b>
AHI (TST) (n/h)	17.4	3.0–81.7	16.3	0–50.1	11.5	0.5–63.8	11.9	0–35.4	<b>0.03</b>
OAH1 (n/h)	0.0	0–0.3	0.1	0–0.2	0.0	0–0.1	0.0	0–0.0	<b>0.039</b>
ODI3 (n/h)	26.5	7.2–79.5	18.8	0–64.0	19.1	0.6–75.5	13.9	0–46.9	0.076

**Table 5**

Mean values from the present and previous studies compared statistically. All bolded values are mean (SD). The *p* values refer the statistical difference between the mean values of the present data and the compared study; statistically significant *p*-values are in italics. The difference between the means; confidence interval of the difference are found in parentheses.

Study	<i>n</i>	IS% Mean (SD)	QS% Mean (SD)	AS% Mean (SD)	IS + AS% Mean (SD)	RDI (/h) Mean (SD)
Present study, all infants	<b>84</b>	<b>21.6 (5.8)</b>	<b>41.5 (5.2)</b>	<b>36.7 (6.9)</b>	<b>58.4 (5.2)</b>	<b>19.3 (3.7)</b>
Anders and Keener (1985)	40	NA	<b>29.0 (7.9)</b> <i>p</i> < 0.0001 (12.6; 10.2–14.9)	<b>51.9 (10.8)</b> <i>p</i> < 0.0001 (–15.2; –18.4 to –12.0)	NA	NA
Hoppenbrouwers et al. (1988)	10	<b>14.4 (5.5)</b> <i>p</i> = 0.0003 (7.3; 3.4–11.1)	<b>35.1 (7.9)</b> <i>p</i> = 0.0007 (6.5; 2.8–10.2)	<b>50.6 (8.2)</b> <i>p</i> < 0.0001 (–13.9; –18.6 to –9.1)	NA	NA
	10	<b>23.6 (8.0)</b> <i>p</i> = 0.3215 (–2.0; –6.1 to 2.0)	<b>29.4 (10.4)</b> <i>p</i> < 0.0001 (12.2; 8.2–16.1)	<b>46.9 (7.4)</b> <i>p</i> < 0.0001 (–10.2; –14.9 to –5.6)	NA	NA
Hoppenbrouwers et al. (2005)	60	<b>4.1 (4.5)</b> <i>p</i> < 0.0001 (17.5; 15.8–19.3)	<b>37.2 (10.3)</b> <i>p</i> = 0.0011 (4.4; 1.8–6.9)	<b>30.2 (11.5)</b> <i>p</i> < 0.0001 (6.5; 3.4–9.6)	NA	NA
Present study, infants <44 weeks	<b>31</b>	<b>22.5 (6.1)</b>	<b>41.2 (4.1)</b>	<b>36.7 (6.2)</b>	<b>59.2 (3.5)</b>	<b>20.7 (3.7)</b>
Montemitro et al. (2008)	10	NA	<b>19 (6)</b> <i>p</i> < 0.0001 (22.2; 18.7–25.6)	NA	<b>59 (11)</b> <i>p</i> = 0.9459 (0.2; –4.3 to 4.6)	<b>16.1 (3)</b> <i>p</i> = 0.001 (4.6; 2.0–7.2)
Coons and Guilleminault (1982)	10	<b>32.8 (6.6)</b> <i>p</i> < 0.0001 (–10.3; –15.0 to –5.7)	<b>31.0 (4.6)</b> <i>p</i> < 0.0001 (10.2; 7.0–13.3)	<b>31.6 (6.3)</b> <i>p</i> = 0.0340 (5.1; 0.4–9.7)	NA	NA
Present study, infants ≥44 weeks	<b>53</b>	<b>21.1 (5.5)</b>	<b>41.8 (5.7)</b>	<b>36.7 (7.3)</b>	<b>57.8 (6.0)</b>	<b>18.4 (3.4)</b>
Coons and Guilleminault (1982)	10	<b>32.0 (7.1)</b> <i>p</i> < 0.0001 (–10.9; –14.9 to –6.9)	<b>37.8 (7.0)</b> <i>p</i> = 0.0543 (4.0; –0.1 to 8.1)	<b>24.7 (4.3)</b> <i>p</i> < 0.0001 (12.0; 7.2–16.8)	NA	NA

RDI = Respiratory Disturbance Index, cortical arousal index (*n*/h of TST).

#### 4. Discussion

The main aim of the present study was to obtain sleep stage and arousal parameters for normal one month old infants by applying the AASM scoring manual as closely as possible. The over-night PSGs were recorded in 84 healthy and full-term one-month-old infants at their homes. There was large individual variability between the infants' sleep parameters in the present study, as according to our data, the  $\pm 2$  SD reference limits for normal over-night sleep in one month old infants were: QS 31–52%, AS 23–51% and IS 10–33%. The individual variability in the amounts of sleep stages is thought to represent an intrinsic characteristic of the maturing central nervous system. It is possible that the young infant's sleep pattern responds unpredictably to various stimuli (e.g. changes in the sleeping environment, or recording equipment), which might at least partly explain the great variability (Hoppenbrouwers et al., 1988; Hoppenbrouwers, 1992).

The acquired sleep parameters were compared to six previous studies (Coons and Guilleminault, 1982; Anders and Keener, 1985; Hoppenbrouwers et al., 1988, 2005; Kirjavainen et al., 2004; Montemitro et al., 2008). Except for the study of Coons and Guilleminault (1982), our scoring method revealed more QS than the others, but the amount of AS or IS were not consistently higher or lower. The following issues could explain the discrepancies across the studies.

In a population of great individual variability, such as infants, a small sample size may not describe the normal range of parameters adequately. The definition of normality as an inclusion criterion in a normative study is equally important, as the variation among the abnormal infant population is thought to be even greater. The discrepancies between the six compared studies and our data raise the question of possible sample selection, especially with the small sample sizes in some of the compared studies. All studies included "healthy and full term" infants, but the definition of normality was variable and the sample size varied from 10 to 84. The number of studied infants (84) was largest in our study. As

specified in the inclusion criteria in the Methods section, our subjects were full term and considered normal and healthy. They had not presented any apparent abnormality while their stay at the maternity hospital or during their routine visits (first two visits usually at 2–3 weeks and 5–6 weeks of age) at the local child health center. In the study by Coons and Guilleminault (1982) and Hoppenbrouwers et al. (2005) no other definition for normality than "healthy and full term" was given, the sample sizes were 10 + 10 and 60 in the two studies, respectively. The study by Anders and Keener (1985) included 40 full term infants, who were recruited from the birth records of the medical center's newborn nursery. The inclusion criteria demanded normal full term status, and normality of pregnancy and birth. The 10 + 10 infants included in the study by Hoppenbrouwers et al. (1988) were considered healthy and selected based on the absence of disease in their mothers. The infants' weights were appropriate for their gestational ages. Kirjavainen et al. (2004) included as a control group 23 full term infants, who had had an uneventful neonatal history. Each infant was also examined by a child neurologist. The sample size in the study by Montemitro et al. (2008) was smallest (10), but their inclusion criteria were detailed. The included infants were born full term, they were born to nonsmoking parents who used no alcohol or drugs, and they were habitually put to sleep in the supine position. They had no family or personal history of apnea or SIDS. At the time of the study the infants were healthy, not sleep deprived, and not medicated.

The choice of scoring criteria could explain some of the variation in sleep stages. As there are no generally accepted sleep scoring criteria for very young infants, the researchers have had to create their own scoring criteria or modify the pre-existing ones to suite the setting at hand. However, even when the same criteria have been applied, the results have differed (Coons and Guilleminault (1982) vs. Montemitro et al. (2008), Hoppenbrouwers et al. (1988) group 1 vs. group 2). Increasing number of state parameters that are required to define a state is known to increase the percentage of IS (Parmelee et al., 1967). Our results, however, do not support this,

as the study by [Hoppenbrouwers et al. \(1988\)](#) had a fixed number of required state parameters, but their data showed one of the lowest amounts of IS. Moreover, the very detailed scoring system in the CHIME (Collaborative Home infant Monitoring Evaluation) criteria applied in [Hoppenbrouwers et al. \(2005\)](#) produced the lowest amount of IS.

Of the six compared studies the study by [Anders and Keener \(1985\)](#) differed from the rest methodologically as it was a time-lapse video study. The remaining five studies, as well as the present study, were PSG studies. All five compared studies included one or two EEG channels and chin EMG. All except for the study by [Coons and Guilleminault \(1982\)](#) also included EOG. All five compared studies also included one or more channels of respiration enabling the evaluation of regularity of breathing. The inclusion of EOG in the set-up should affect the scoring of AS, theoretically, but the studies without EOG ([Anders and Keener, 1985](#); [Coons and Guilleminault, 1982](#)) did not show systematically more or less AS than the rest.

We used six EEG channels in sleep staging, which may improve the scoring accuracy as compared to the other five that utilized one or two EEG derivations. In theory, the accuracy might benefit from multiple EEG-channels because of topographical differences in sleep EEG. The Anders manual ([Anders et al., 1971](#)) states, that SWA is maximal over frontal regions in younger children. One topographical study on the newborns' EEG showed the increase of slow activity mainly over central and temporal electrode locations with increasing age ([Pereda et al., 2006](#)). Therefore, the number of recorded EEG channels is not likely a significant factor in the inter-study differences in infants. The nasal cannula was left out of the set-up in our study to avoid unnecessarily disturbing normal sleep. This does not, however, explain the higher amount of QS in our study, as none of the six compared studies had a nasal pressure transducer, either. Still, when asked to compare the PSG night to the infant's normal night, the parents in our study often admitted a difference: in total, 56% of the parents felt, that the night had been worse or better than a normal night. However, the parent-reported quality of sleep did not reflect to the sleep parameters; the sleep stage proportions and arousal and awakening indices did not statistically differ between the infants with better and worse parent-reported sleep quality. The infants seem to react to study conditions differently. On the other hand, the common notion is that infants do not express significant first-night effect ([Coons and Guilleminault, 1982](#)) and their night-to-night variability in sleep stages or apneas is minimal ([Rebuffat et al., 1994](#)). This might indicate that the individual direction of reaction is constant and the recording set-up does not significantly influence the sleep parameters in infants.

Most of the previously used scoring criteria include observational information in the decision making. Observation (annotations or video) was included in four of the six studies that were compared to our study ([Anders and Keener, 1985](#); [Hoppenbrouwers et al., 1988, 2005](#); [Montemitro et al., 2008](#)), while the present study, [Kirjavainen et al. \(2004\)](#) and the study by [Coons and Guilleminault \(1982\)](#) did not have continuous observation. With observational data, one could assume that especially the scoring of quiet wakefulness would be more accurate and result in less IS and AS; the amount of AS would be overestimated without knowing if the eyes are open or closed. Also the differentiation of drowsy wakefulness from IS would sharpen. The three studies with no observational data did not, however, find consistently different amounts of IS than the rest. As a scoring method, observation requires continuous technician resource or a lot of time-consuming video-playback during the scoring. This advocates the development of reliable visual scoring rules based on EEG and supplementary polysomnographic data.

The length of a scoring epoch is known to influence the sleep parameters. The amount of AS increases and QS decreases when

the epoch lengthens ([Kulp et al., 2000](#)): a longer epoch is more likely to include movement or eye movements. The epoch length was 30 s in the present study, in [Hoppenbrouwers et al. \(2005\)](#), [Coons and Guilleminault \(1982\)](#), [Kirjavainen et al. \(2004\)](#), and in [Montemitro et al. \(2008\)](#). In [Hoppenbrouwers et al. \(1988\)](#) it was one minute and in [Anders and Keener \(1985\)](#) it was five minutes. Indeed, the studies by [Hoppenbrouwers et al. \(1988\)](#) and [Anders and Keener \(1985\)](#) reported more AS than the rest. The length of the smoothing window does not have a significant effect on sleep stage percentages ([Kulp et al., 2000](#)).

The experience of the scoring personnel may also play a role, and concerning normative values, the inter-scorer agreements of visual sleep scoring should be presented. Our PSGs were scored by two clinical neurophysiologists with experience in the visual sleep scoring and infant EEG. Among the six compared studies, only [Hoppenbrouwers et al. \(1988, 2005\)](#) acknowledge the scoring personnel as "10 different scorers" and "trained personnel", respectively. The other studies did not identify the scorers. Moreover, the reproducibility of the hypnograms should be taken into account. The young infants' immature state differentiation imposes some ambiguity on visual scoring. IS, by definition, does not fulfill the criteria for any one stage. The age interval between the disappearance of tracé alternant pattern and the emergence of sleep spindles (about 4–6 weeks to 2–3 months) makes also QS vulnerable to inter-scorer disagreement. In young infants, there is a known lack of concordance between the EMG tone, respiratory pattern and other PSG signals ([Dreyfus-Brisac, 1970](#); [Parmelee et al., 1972](#); [Schloon et al., 1976](#)). In our work the inter-scorer agreement yielded a substantial overall agreement (80.6%), which is a good achievement considering the challenging age of the infants. The agreement in QS, AS, and W were more than 80%, while as expected, the agreement in IS was lower, 46%. Of the six compared studies only [Anders and Keener \(1985\)](#), [Hoppenbrouwers et al. \(1988, 2005\)](#) define the methods to ensure reproducibility. In the [Anders and Keener \(1985\)](#) study 20% of the nights and a two hour sample of each night were randomly selected and subjected to double scoring. This produced an inter-rater concordance of 0.76–0.99 and median kappa scores for QS, AS, and W 0.89, 0.83, and 0.89, respectively. In the [Hoppenbrouwers et al. \(1988\)](#) study total of 10 people functioned as scorers over the five-year period. The scorers were trained in the applied scoring rules until the inter-scorer agreement was >80%. Moreover, a 20 min sample of each recording was double scored, and if the agreement was <80%, the entire recording was rescored. This resulted in >80% agreement in QS, AS and W, but only about 50% in IS. This is in line with our results. As a part of the multi center CHIME study, the recordings in [Hoppenbrouwers et al. \(2005\)](#) were scored at the CHIME analysis center with trained personnel, whose reliability of the scoring has been evaluated and published earlier ([Crowell et al., 1997](#)).

Usually the PSGs of young infants are performed at hospital wards or intensive care units. Our recordings were made at home, under as neutral conditions as possible since home recordings have been recommended to give a realistic picture of infants' sleep ([Bernstein et al., 1973](#)). On the other hand, the home environment is not as exactly controlled as in laboratory. In the light of the six previous studies we used for comparison, the studies with home recordings have not revealed parameters consistently different from those recorded in laboratory. Neither can the infants' individual variability be attributed to the study setting or the burden of the recording equipment, because the range of sleep parameters has been wide irrespective of these conditions. We therefore believe that the recording place has little influence on the reference values in the young infants.

The effect of recording length (over-night vs. 24 h) does not seem a plausible factor for sleep parameter differences, either,



because the circadian regulation starts to show at about 5–6 weeks of age and a clear diurnal pattern with distinct daytime naps can be seen by 12 weeks of age (Coons and Guilleminault, 1982). Homeostasis-related temporal organization of sleep states begins to emerge in the second month of life. (Parmelee et al., 1964; Schechtman et al., 1994; Coons and Guilleminault, 1982; Hoppenbrouwers et al., 1982) Indeed, the studies including also the day-time in their parameters, do not report consistently different percentages of sleep stages in TST/24 h. To sum up the above discussion on the possible factors behind the inter-study differences, no obvious explanations stand out.

In addition to the amounts of sleep stages, the cortical arousals were also quantified in the present study. The current scoring criteria for arousals (Berry et al., 2012) are not easily applicable for young children and infants. This is because according to the rules, the EEG frequency shift in the delta range during an arousal does not qualify. However, infant arousals are known to often consist of delta activity. Normal dominant posterior activity, after its emergence at about 3–4 months of age, is in the delta range, and, even before the rhythmical and organized posterior activity, the waking EEG activity in an infant consists of irregular, medium-to high voltage delta activity. (Lindsley, 1936; Dreyfus-Brisac, 1975). The IPWG consensus (The IPWG, 2005) recommended scoring delta arousals in young children and infants, but the Pediatric Task Force, during the preparation of AASM 2007 scoring manual (Grigg-Damberger et al., 2007; Iber et al., 2007), decided not to. The task force could not agree whether delta arousals in children represent true arousals. Moreover, the task force members were concerned that it might be difficult to distinguish delta arousals from slow wave sleep. IPWG also recommended scoring subcortical activations, i.e. arousal events with no visible change in the EEG, but AASM task force decided not to endorse scoring these, because their clinical significance is unclear. The normal values concerning cortical arousals in infants are rarely presented. In our data the awakening index was 4–13/h and cortical arousal index 12–27/h. Of the six compared studies only the one by Montemitro et al. (2008) present the arousal indices for ten three-week-old subjects. Compared to that study, our arousal index was higher. Some of this difference might result from the six-channelled EEG-montage in our study instead of two-lead EEG. More EEG-channels is known to result in higher arousal indices (O'Malley et al., 2003).

We found a statistically significant correlation between the percentage of IS and the awakening index; the infants with more awakenings presented a greater percentage of IS. However, the other correlations between the arousal and awakening indices with the sleep stage percentages were not significant statistically.

The age of the infants we studied was on average 44.7 weeks at the recording (range 42.3–48.4 weeks). Thirty-one infants (37%) were less than 44 weeks of age at the recording. We compared the sleep parameters in the two age groups, the infants under 44 weeks and the infants 44 weeks or older. The sleep stage percentages in our data were not statistically different between the groups, although in previous studies the proportion of AS has been found to be decreasing and QS increasing during early life (Hoppenbrouwers et al., 1982, 1988). However, sleep was more continuous in the older age group, assessed by the decreased number of arousals and awakenings, which is consistent with the previous studies (Anders and Keener, 1985; Montemitro et al., 2008) and can be considered an effect of maturation.

## 5. Conclusions

The large overlap in the sleep parameters of normal and abnormal infants is a recognized problem in the clinical work

(Hoppenbrouwers et al., 1988). Because of the large individual variation in sleep quality, only large deviations from the average can stand alone as a marker of abnormality. The variability between infants also emphasize the importance of reliable and age-appropriate reference values for sleep parameters in the diagnostics of sleep disorders. Therefore, many studies have aimed to establish these values over the past decades. However, the resulting reference values have been difficult to put to general use, because the scoring criteria and the methods by which the values have been acquired have varied from study to study. Uniform sleep scoring methods and criteria might make the normative values in the literature more applicable to the clinical work. The ideal solution would be that the same rules and methods could be applied to all patients regardless of the age.

Even though the scoring rules for adults and children older than two months of age have been revised recently (Berry et al., 2012), there still are no universally accepted scoring rules for infants younger than two months of age. We adapted the AASM rules for this age group. The rules turned out to be well applicable and reproducible for the unattended PSGs in one-month-olds. The independence from visual observation or video playback can be considered an advantage of these rules over many other scoring manuals.

In infants, the PSG parameters must always be evaluated against the backdrop of accurate sleep–wake scoring because the control of breathing and other physiological phenomena depend on the sleep stage. While there is much overlap in sleep quality in the normal and abnormal infants, in some PSG parameters, there is known to be very little variation in healthy infants. For example, the incidence of obstructive events is known to be uniformly very low in normal infants, which makes even small differences count. To address the problem of wide variation of normality, the quantitative analyses, especially the EEG topography, could give supplementary information about the normality of an infant's sleep. As the early infancy is the period of intensive sleep development, quantitative measures might also reveal subtle differences between, for example, groups of only slight age differences.

## Acknowledgements

The study was funded by Academy of Finland, Grant no. 134880 and by the Competitive Research Financing of the Expert Responsibility area of Tampere University Hospital, Grant nos. 9R007 and 9P013. AS has received grants from Finnish Sleep Research Society and The Research Foundation of the Pulmonary Diseases.

We thank Tytti Koskelo, who organized and performed the recordings. We are also grateful to Anneli Kylliäinen for reviewing this manuscript.

*Conflict of interest:* None of the authors have potential conflicts of interest to be disclosed.

## References

- Anders T, Emde R, Parmelee A. A manual of standardized terminology, techniques and criteria for scoring states of sleep and wakefulness in newborn infants. Los Angeles: UCLA Brain Information Service, NINDS Neurological Information Network; 1971.
- Anders TF, Sostek AM. The use of time lapse video recording of sleep–wake behavior in human infants. *Psychophysiology* 1976;13:155–8.
- Anders T, Keener M. Developmental course of nighttime sleep–wake patterns in full-term and premature infants during the first year of life. I. *Sleep* 1985;8:173–92.
- Aserinsky E, Kleitman N. A motility cycle in sleeping infants as manifested by ocular and gross bodily activity. *J Appl Physiol* 1955;8:11–3.
- Bernstein P, Emde R, Campos J. REM sleep in four-month infants under home and laboratory conditions. *Psychosom Med* 1973;35:322–9.
- Berry RB, Budhiraja R, Gottlieb DJ, Gozal D, Iber C, Kapur VK, Marcus CL, Mehra R, Parthasarathy S, Quan SF, Redline S, Strohl KP, Davidson Ward SL, Tangredi MM. Rules for scoring respiratory events in sleep: update of the 2007 AASM manual for the scoring of sleep and associated events. *Deliberations of the sleep apnea*

- definitions task force of the American Academy of Sleep Medicine. *J Clin Sleep Med* 2012;8:597–619.
- Coons S, Guilleminault C. Development of sleep–wake patterns and non-rapid eye movement sleep states during the first six months of life in normal infants. *Pediatrics* 1982;69:793–8.
- Crowell D, Brooks L, Colton T, Corwin M, Hoppenbrouwers T, Hunt C, Kapuniia L, Lister G, Neuman M, Peucker M, Davidson Ward S, Weese-Meyer D, Willinger M. CHIME steering committee. Infant polysomnography: reliability. *Sleep* 1997;20:553–60.
- Dement W, Kleitman N. Cyclic variations in EEG during sleep and their relation to eye movements, body motility, and dreaming. *Electroencephalogr Clin Neurophysiol* 1957;9(Suppl.):673–90.
- Dittrichová J. Development of sleep in infancy. *J Appl Physiol* 1966;21:1243–6.
- Dreyfus-Brisac C. Ontogenesis of sleep in human prematures after 32 weeks of conceptual age. *Dev Psychobiol* 1970;3:91–121.
- Dreyfus-Brisac C. Neurophysiological studies in human premature and full-term newborns. *Biol Psychiatry* 1975;10:485–96.
- Gertner S, Greenbaum CW, Sadeh A, Dolfin Z, Sirota L, Ben-Nun Y. Sleep–wake patterns in preterm infants and 6 month's home environment: implications for early cognitive development. *Early Hum Dev* 2002;68:93–102.
- Goodwin JL, Enright PL, Kaemingk KL, Rosen GM, Morgan WJ, Fregosi RF, Quan SF. Feasibility of using unattended polysomnography in children for research – report of the Tucson children's assessment of sleep apnea study (TuCASA). *Sleep* 2001;24:937–44.
- Grigg-Damberger M, Gozal D, Marcus CL, Quan SF, Rosen CL, Chervin RD, Wise M, Picchietti DL, Sheldon SH, Iber C. The visual scoring of sleep and arousal in infants and children. *J Clin Sleep Med* 2007;3:201–40.
- Guilleminault C, Souquet M. Scoring criteria. In: Guilleminault C, editor. *Sleeping and waking problems: indications and techniques*. Menlo Park, CA: Addison-Wesley Publishing Co.; 1982. p. 415–26.
- Haddad GG, Jung HJ, Lai TL, Mellins RB. Determination of sleep state in infants using respiratory variability. *Pediatr Res* 1987;21:556–62.
- Halbower AC, Degaonkar M, Barley PB, Earley CJ, Marcus CL, Smith PL, Prahme MC, Mahone EM. Childhood obstructive sleep apnea associates with neuropsychological deficits and neuronal brain injury. *PLoS Med* 2006;3:e301.
- Hellström-Westas L, de Vries L, Rosen I. Atlas of amplitude-integrated EEG's in the newborn. Boca Raton (FL): Parthenon Publishing Group; 2003.
- Hoppenbrouwers T, Hodgman JE, Harper RM, Sterman MB. Temporal distribution of sleep states, somatic activity during the first half year of life. *Sleep* 1982;5:131–44.
- Hoppenbrouwers T. *Sleep in infants*. New York: Raven Press; 1987.
- Hoppenbrouwers T, Hodgman J, Arakawa K, Geidel SA, Sterman MB. Sleep and waking states in infancy: normative studies. *Sleep* 1988;11:387–401.
- Hoppenbrouwers T. Polysomnography in newborns and young infants: sleep architecture. *J Clin Neurophysiol* 1992;9:32–47.
- Hoppenbrouwers T, Hodgman JE, Rybina D, Fabrikant G, Corwin M, Crowell D, Weese-Meyer DE. CHIME study group. Sleep architecture in term and preterm infants beyond the neonatal period: the influence of gestational age, steroids, and ventilator support. *Sleep* 2005;28:1428–36.
- Iber C, Ancoli-Israel S, Chesson A, Quan S. For the American Academy of Sleep Medicine. *The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications*. 1st ed. Westchester, IL: American Academy of Sleep Medicine; 2007.
- Jenni OG, Borbély AA, Achermann P. Development of the nocturnal sleep electroencephalogram in human infants. *Am J Physiol Regul Integr Comp Physiol* 2004;286:R528–38.
- Kirjavainen J, Lehtonen L, Kirjavainen T, Kero P. Sleep of excessively crying infants: a 24-hour ambulatory sleep polygraphy study. *Pediatrics* 2004;114:592–600.
- Kulp TD, Corwin MJ, Brooks LJ, Peucker M, Fabrikant G, Crowell DH, Hoppenbrouwers TCHIME Study Group. The effect of epoch length and smoothing on infant sleep and waking state architecture for term infants at 42–46 weeks postconceptional age. *Sleep* 2000;7:1–7.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.
- Lindsley D. Brain potentials in children and adults. *Science* 1936;84:354.
- Louis J, Zhang J, Revol M, Debilly G, Challamel MJ. Ontogenesis of nocturnal organization of sleep spindles: a longitudinal study during the first six months of life. *Electroencephalogr Clin Neurophysiol* 1992;83:289–92.
- McNamara F, Issa FG, Sullivan CE. Arousal pattern following central and obstructive breathing abnormalities in infants and children. *J Appl Physiol* 1996;81:2651–7.
- Montemirro E, Franco P, Scaillet S, Kato I, Groswasser J, Villa MP, Kahn A, Sastre JP, Ecohard R, Thiriez G, Lin JS. Maturation of spontaneous arousals in healthy infants. *Sleep* 2008;31:47–54.
- O'Brien LM, Gozal D. Behavioural and neurocognitive implications of snoring and obstructive sleep apnea in children: facts and theory. *Paediatr Respir Rev* 2002;3:3–9.
- O'Malley EB, Norman RG, Farkas D, Rapaport DM, Walsleben JA. The addition of frontal EEG leads improves detection of cortical arousal following obstructive respiratory events. *Sleep* 2003;26:435–9.
- Parmelee AH, Wenner W, Akiyama Y, Schulz M, Stern E. Sleep states in premature infants. *Dev Med Child Neurol* 1967;9:70–7.
- Parmelee AH, Stern E. Development of states in infants. In: Clemente CD, Purpura DP, Mayer FE, editors. *Sleep and the maturing nervous system*. New York: Academic Press; 1972. p. 199–228.
- Parmelee AH, Stern E, Harris MA. Maturation of respiration in preterm and young infants. *Neuropadiatr* 1972;3:294–304.
- Parmelee AH, Wenner W, Schulz H. Infant sleep patterns from birth to 16 weeks of age. *J Pediatr* 1964;65:576–82.
- Pereda E, de La Cruz D, Mañas S, Garrido J, López S, González J. Topography of EEG complexity in human neonates: effect of the postmenstrual age and the sleep state. *Neurosci Lett* 2006;394:152–7.
- Rebuffat E, Groswasser J, Kelmanson I, Sottiaux M, Kahn A. Polygraphic evaluation of night to night variability in sleep characteristics and apnea in infants. *Sleep* 1994;17:329–32.
- Sankupellay M, Wilson S, Heussler H, Parsley C, Yuill M, Dakin C. Characteristics of sleep EEG power spectra in healthy infants in the first two years of life. *Clin Neurophysiol* 2011;122:236–43.
- Schechtman V, Harper R, Harper R. Distribution of slow-wave activity across the night in developing infants. *Sleep* 1994;17:316–22.
- Schloon H, O'Brien MJ, Scholten CA, Precht HF. Muscle activity and postural behaviour in newborn infants. A polymyographic study. *Neuropadiatr* 1976;7:384–415.
- So K, Adamson TM, Horne RSC. The use of actigraphy for assessment of the development of sleep/wake patterns in infants during the first 12 months of life. *J Sleep Res* 2007;16:181–7.
- The International Pediatric Work Group on Arousals (IPWG). The scoring of arousals in healthy term infants (between the ages of 1 and 6 months). *J Sleep Res* 2005;14:37–41.
- Tikotzky L, De Marcas G, Har-Toov J, Dollberg S, Bar-Haim Y, Sadeh A. Sleep and physical growth in infants during the first 6 months. *J Sleep Res* 2010;19:103–10.