Sleep Medicine 101 (2023) 260–268

Contents lists available at ScienceDirect

Sleep Medicine

journal homepage: www.elsevier.com/locate/sleep

Sleep assessment in preterm infants: Use of actigraphy and aEEG

Özge Ülgen [a], Hatice Ezgi Barış [a], Öykü Özbörü Aşkan [a], Selda Küçük Akdere [a], Can Ilgın [b], Hülya Özdemir [c], Nural Bekiroğlu [d], Kıvılcım Gücüyener [e], Eren Özek [c], Perran Boran [a, *]

[a] Marmara University, School of Medicine, Department of Pediatrics, Division of Social Pediatrics, Istanbul, Turkey
[b] Marmara University, School of Medicine, Division of Public Health, Istanbul, Turkey
[c] Marmara University, School of Medicine, Department of Pediatrics, Division of Neonatology, Istanbul, Turkey
[d] Marmara University, School of Medicine, Department of Biostatistics, Istanbul, Turkey
[e] Gazi University, School of Medicine, Division of Pediatric Neurology, Ankara, Turkey

ARTICLE INFO

Article history:
Received 22 October 2022
Received in revised form
18 November 2022
Accepted 18 November 2022
Available online 20 November 2022

Keywords:
Actigraphy
Amplitude-integrated electroencephalography
Preterm
Newborn
Sleep-wake cycle
Development

ABSTRACT

Objective: Objective methods to monitor the sleep of preterm infants at the neonatal intensive care unit (NICU) are required to prevent potentially adverse neurodevelopmental outcomes. This study aimed to determine the concordance of actigraphy and amplitude-integrated electroencephalogram (aEEG) against gold standard direct observation (DO) in assessing sleep/wake states of typically developing preterm infants.

Methods: This prospective observational study was conducted in a single center level III NICU. Sleep variables were measured using Philips Respironics Mini-Mitter® Actiwatch-2 for 24 h and compared with 8-h matched data of aEEG and DO. Sensitivity-specificity analysis, Cohen's kappa, prevalence-adjusted and bias-adjusted kappa (PABAK), and Bland Altman plots were generated.

Results: Seventeen preterm infants were recruited. A total of 11252 epochs were studied. Sensitivity (86.4%), agreement rate (67.9%), and predictive value for wake (47.9%) for the actigraphy were highest at the automatic activity threshold whereas specificity (54.5%) and predictive value for sleep (75.5%) were highest at low threshold. The sensitivity of aEEG was 79.3% and the specificity was 54.3%. At all thresholds, the agreement was largely equivalent with low kappas (0.14–0.17) and PABAK coefficients (0.22–0.35) for actigraphy and DO. Moderate agreement was observed between aEEG and DO according to the PABAK coefficient (0.44). Mean differences in sleep parameters were not different between DO and aEEG as well as DO/aEEG and actigraphy at medium threshold ($p > 0.05$).

Conclusions: Actigraphy at medium threshold can be used in depicting sleep in typically developing preterm infants at NICU. aEEG may be an alternative adjunctive method to actigraphy for the evaluation of sleep/wake states in the NICU setting.

Clinical trial registration number: NCT04145362

1. Introduction

The estimated global preterm birth rate is 10.6% of all live births and this number tends to increase each year [1]. Consistent with the global rate, the reported preterm birth rate in Turkey is 11% [2]. Advanced care in neonatology resulted in decreased mortality, but morbidity related to prematurity remained an important concern [3]. Sequelae of prematurity include adverse short- and long-term neurodevelopmental outcomes such as severe sensory or motor disabilities, cognitive deficits, social and emotional problems as well as sleep problems [3,4].

The development of sleep-wake behavior is an important maturation process that requires functional connectivity of the neural network starting from fetal life and continuing particularly during the first months after birth [5]. Impaired sleep has been shown to disturb the myelination of the maturing brain [6]. Preterm infants are at significant risk for altered sleep development. Due to prematurity, the establishment of the neural connectivity framework is interrupted or immature which interferes with sleep development. In addition, the preterm brain is prone to multiple insults such as hypoxia, hyperoxia, or inflammation resulting in

* Corresponding author. Muhsin Yazıcıoğlu Street, No:10, 34899, Pendik/İstanbul, Turkey.
E-mail addresses: drperran@yahoo.com, drperran@gmail.com (P. Boran).

https://doi.org/10.1016/j.sleep.2022.11.020
1389-9457/

Abbreviations

AAP	American Academy of Pediatrics
ABSS	Anderson Behavioral State Scale
aEEG	Amplitude-integrated electroencephalogram
AR	Agreement rate
BSID-III	Bayley Scales of Infant Development
CA	Chronological age
CS	cramp-synchronized
DO	direct behavioral observation
GA	Gestational age
GM	General Movement
NICU	Neonatal Intensive Care Unit
NTISS	Neonatal Therapeutic Intervention Scoring System
PABAK	prevalence-adjusted and bias-adjusted kappa
PMA	Postmenstrual age
PR	Poor-Repertoire
PSG	Polysomnography
PVS	predictive value of sleep
PVW	predictive value of wake
WASO	Wake after sleep onset

injury. Instead of the protective regulatory in utero environment, the physical factors (light, temperature, sound) and frequent interventions may disturb the sleep periods of preterm infants in the Neonatal Intensive Care Unit (NICU) [7,8]. High lighting levels and loud sound levels have been associated with sleep disturbances among other adverse clinical outcomes [9].

Significant research showed the persisting effects of prematurity on sleep throughout childhood [5]. Sleep problems were more frequently observed in 11-year-old children born preterm regardless of the presence of neurodevelopmental disabilities [10]. Shorter gestational age at birth was associated with longer sleep duration in the early years of life in three population-based cohorts [11]. In contrast, 3–5 years old preterm born children were more likely to have shorter sleep duration and adverse sleep outcomes [12]. Shorter nighttime sleep duration in preterm children was associated with social and emotional problems [13].

Investigating the sleep development of preterm infants in NICU provides a unique opportunity to understand the potential implications of disrupted sleep on early neurodevelopment. However, monitoring sleep in fragile preterm infants is challenging. Available reliable methods such as direct behavioral observation (DO) and polysomnography (PSG) are obtrusive and require expertise to interpret or technical support. Unobtrusive methods to measure infant sleep with less burden on vulnerable preterm infants and caregivers in NICU are necessary to incorporate sleep monitoring into routine neonatal care of preterm infants [14].

Actigraphy is a non-invasive tool to estimate sleep-wake cycles based on activity monitoring, that has been increasingly used in pediatric sleep research [15]. It allows objective sleep assessment in an infant's natural environment and is recognized as a valid and useful tool for sleep evaluation in children. Few studies investigated the use of actigraphy in preterm children and revealed divergent results due to the differences in methodology and study participants [16–21]. Amplitude-integrated electroencephalogram (aEEG) is another readily available bedside tool to monitor cerebral function. It has widespread use to predict the neurodevelopmental outcome of term and preterm infants in the NICU [22]. Neonatal sleep-wake cycles can also be identified by aEEG [23]. In this study, we aimed to determine the concordance of actigraphy and aEEG against DO in assessing the sleep/wake states of preterm babies.

2. Methods

2.1. Participants and settings

This prospective observational study was conducted in a single center Level III NICU at Marmara University Hospital located in Istanbul, Turkey between June 2019 to May 2020. The study was approved by the Marmara University Ethics Committee (protocol ID: 09.2017.488) and registered with the Clinical Trials (NCT04145362).

Eligible preterm infants were recruited by convenience sampling. All preterm infants born at gestational age (GA) 28–37 weeks and admitted to the unit during the study period (n = 70) were eligible. Fifty-three infants were excluded according to the exclusion criteria (Fig. 1). Parents of all consecutively admitted neonates who met inclusion criteria were invited to participate in the study. Written informed consent was obtained from both parents of each infant. Infants whose parents refused to participate in the study did not differ concerning gender, GA, or birth weight.

Routine medical care was provided according to the NICU procedures, and the unit routine was not changed during the study period. Infants were fed every 2–3 h and exposed to cyclic lighting daily.

2.2. Procedures and measurements

Demographic and clinical data were collected from the patients' medical records. Gestational age was assessed by the last menstrual period and/or by prenatal ultrasonography and was confirmed by Ballard Scoring System. Postmenstrual age (PMA) was determined as GA plus the time since birth (weeks). Chronological age (CA) was defined as the time since birth (days) [24]. Moderate to late preterm infants were infants born at GA between 32 weeks and $36^{6/7}$ weeks and very preterm infants were born at GA between 28 weeks and 32 weeks [25]. Infants' illness severity was measured by the Neonatal Therapeutic Intervention Scoring System (NTISS) which is accepted as a validated direct measure of resource utilization associated with mortality risk estimates [26]. The higher scores indicated higher disease severity. NTISS for each patient was scored at admission to the NICU and at the moment of the study measurements.

2.2.1. Assessment of sleep

Infants with PMA >32 weeks were found to have more distinct sleep/wake states and better recognizable sleep architecture [27]. Therefore, measures were conducted after 32 weeks of PMA as soon as the babies were clinically stable and under no sedative medication.

2.2.1.1. Actigraphy. Sleep variables were measured using Philips Respironics Mini-Mitter® Actiwatch-2 for 24 h. Actiwatch-2 devices (weighing 16 g) were placed on the right calf (midpoint between the knee and the ankle) except for four babies due to intravenous catheter. Actiware 6.0.9 software was used to calculate sleep-wake patterns. The Epoch length studied was 30 s [21]. The scoring algorithm and wake threshold sensitivity levels were based on previously published validation studies in preterm infants [13–15,19]. Sleep onset was defined as 5 consecutive minutes of decreased activity scored as sleep. Data were analyzed at each of the low (activity value 20), medium (activity value 40), high (activity value 80), and automatic (mean activity counts*0.888/epoch length) activity threshold settings. The studied period was coded as sleep when the measured activity signal was less than or equal to the activity threshold setting. The number of minutes required to score a wake was >5 min of a wake. The number of minutes

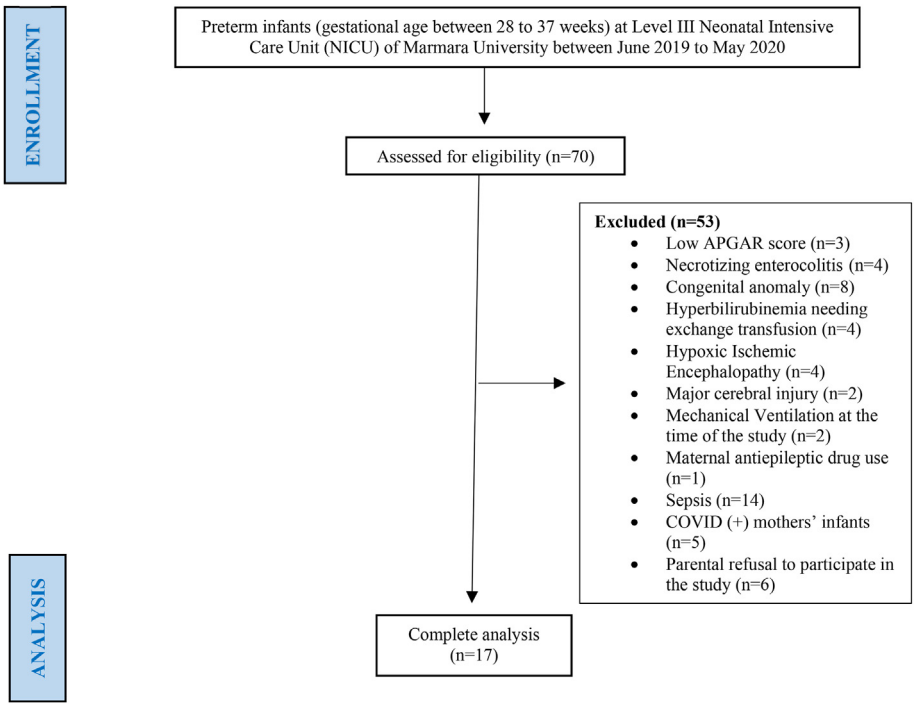

Fig. 1. Flow diagram of the study.

required to score nocturnal wake frequency was >5 min of wake as recommended in previous studies. After each recording, the data were downloaded and stored by the research assistant (ÖÜ) on a personal computer designated for the research. Actigraphy data measurements included total sleep time, sleep efficiency, frequency, and average duration of wake and sleep bouts. Sleep efficiency was defined as total sleep time/total sleep time + total wake time*100. Wake after sleep onset (WASO) was defined as the number of waking minutes between sleep start and end times.

Daytime was defined from 8 a.m. to 8 p.m. and nighttime from 8 p.m. to 8 a.m. based on the NICU routines [28].

Sleep, activity, and feeding diaries were completed by the primary nurse responsible for that infant over 24 h divided into 5 min blocks. Sleep-wake periods, type of clinical interventions, and adverse events were recorded. The NICU practices minimal handling procedures and since the study included only clinically stable newborns, no adverse events during the study period were recorded. Other recorded interventions included diaper changes every 4 h and feeding every 3 h, performed day and night in all infants. Since these procedures were performed both day and night, periods of feeding and care were not excluded from the analysis. The actigraphy data were scored in 30-sec epochs by two independent pediatricians (ÖÜ, PB). Inter-observer agreement was 81% and the kappa score was 0.75 indicating substantial agreement.

2.2.1.2. Amplitude-integrated electroencephalogram. Overall, 4 infants had to be excluded from analysis because of technical failure of aEEG recording, and recordings of 13 infants were used for aEEG. Continuous aEEG recordings were performed after 32nd weeks PMA, when the babies wore the actigraph, using the Cerebral Function Monitor Olympic Brainz Monitor (OBM- Natus Medical Inc, San Carlos, USA). The aEEG was recorded as a four-channel EEG using needle electrodes. Electrodes were placed on the scalp corresponding to the positions C3, C4, and P3, P4 of the international 10–20 system with a reference electrode on the back. The data was retrieved by using OBMViewer 3.1.4.149 software. Recordings were continued for 24 h and synchronized with actigraphy. aEEG scores were assigned according to the Burdjalov scoring system by an experienced pediatric neurologist blinded for the patient's clinical outcome, and the actigraphy data (KG) [29]. Four variables of aEEG recordings were scored: Continuity, sleep-wake cycling (SWC), amplitude of lower border (LBA), and bandwidth span (BW). Individual component scores were summed to determine the total score for each recording. The minimum possible total score was 0, and the maximum was 13.

2.2.1.3. Direct behavioral observation. The infants were video recorded simultaneously as the aEEG and actigraphy recording. The video recording system consisted of an FDR-AX33 visible light camera (Sony Electronics) attached to a tripod. The camera was installed at a height of approximately 80 cm above the infants' incubators to capture the whole bed in view. The 8 h of video recording was selected for direct observation, including 2-h interfeeding periods [4 h of day-time (8.00 a.m.–10 a.m. feeding period, 15.00–17:00 a.m. postfeeding period) and 4 h of night-time measurement (8:00 p.m.-12:00 p.m.)]. The direct observational assessment was recorded using behavioral states from videotaped considering eye states, vocalizations, respiration, and movements. The Anderson Behavioral State Scale (ABSS) which was adapted by Anderson for premature infants from the scale of Parmalee and Stern (1972) was used for the behavioral state evaluation as described in the study of Gill NE et al. [30,31]. The original 12-item scale was used and the dominant (≥16 s) behavioral state was coded for each 30-sec epoch. The behavioral states were defined as follows: 1 = very quiet sleep, 2 = quiet sleep, 3 = restless sleep, 4 = very restless sleep, 5 = drowsy, 6 = alert inactivity, 7 = quiet awake, 8 = restless awake, 9 = very restless awake, 10 = fussing, 11 = crying, 12 = hard crying [31]. States 1–4 were considered as sleep and states 5–12 as wake. A one-week program was given to train the researcher to assess preterm babies with ABSS. Assessment techniques and coding methods were demonstrated with videos in addition to a lecture provided by an experienced

neonatologist and pediatric neurologist. The direct observational assessment was coded by the two researchers (KG, ÖÜ). A pilot of 3 infants was coded separately and intra-rater agreement was 91%.

2.2.2. Assessment of NICU settings

2.2.2.1. Light monitoring. Light exposure was recorded by the Actiwatch light sensor, reflecting the light reaching the infant. Light exposure data in actiwatch devices were downloaded to the computer using Actiware software. Data from one infant is excluded from the analysis since he was under phototherapy at the time of sleep assessment.

2.2.2.2. Sound monitoring. A sound meter was placed in each crib near the head of the infant within 30 cm of the infant's ear to monitor sound exposure for 24 h during the study period. Sound levels were measured by a digital sound level meter (Verth CS 122L China). Version 4.0.3.0 Sound Level Meter Software (SE322) as the data logger. Equivalent sound level (Leq) in A-weighted dB (dBA) was recorded at 3-s intervals.

2.2.3. Neurodevelopmental assessments

A clinical and neurodevelopmental examination was performed at two time points: prior to discharge and at 12 weeks corrected age including General Movement Assessment and Bayley Scales of Infant Development III.

2.2.3.1. General movement assessment. Each infant was video-recorded from above, lying in a supine position, to assess spontaneous movements during active wakefulness without pacifier use. The video recordings lasted for 3–5 min and contained at least three general movement (GM) sequences [32]. A pediatric neurologist trained and experienced in Prechtl's GM assessment analyzed the recordings. Scoring included global judgment as normal, poor-repertoire (PR), cramp-synchronized (CS) and chaotic GMs [32,33]. Fidgety movements were assessed at 12 weeks corrected age.

2.2.3.2. Bayley Scales of Infant Development (BSID-III). Neurodevelopmental outcome was measured by the Bayley Scales of Infant Development (BSID-III) at 12 weeks corrected age. It is a standardized development tool that assesses: Cognitive, communication, fine motor, gross motor, and social/emotional areas [34]. Composite scores from raw scores for all subsets were calculated. Composite scores were scaled to a metric with a mean of 100, SD of 15, and range of 40–160. Composite or composite score equivalent scores >85 were considered as within the normal range, scores 70 to 85 were considered as having a mild learning difficulty, scores between 50 and 70 as moderate learning difficulty, and scores <55 severe learning difficulty in that test area. Bayley-3 standardized scores of <70 (2SD below the normative mean) were investigated as this is the cut-off threshold used to define severe disability in large neuroprotection trials [34]. A developmental specialist who has received professional training in educational and psychological assessment performed and interpreted the test and was blinded to the clinical data of the infant (SKA).

2.3. Statistical analyses

Analyses were conducted using Stata version 15.1 (Stata Corp. 2017. Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC) and IBM SPSS Statistics software (version 28.0, IBM Inc, United States). Differences in continuous variables between the groups (such as GA groups or feeding method) were examined either by Mann-Whitney U or independent samples *t*-test as appropriate and were expressed as mean ± standard deviation (SD) to compare with the previous sleep studies. Day and night sound and light exposure were compared by Wilcoxon signed rank test. Correlation analyses were performed by Spearman's rank correlation. Continuous sleep parameters such as total/nighttime/daytime sleep and awake duration of each subject were compared between each activity thresholds of actigraphy and DO/aEEG using one-way ANOVA repeated measures test with a Greenhouse-Geisser correction and posthoc analysis with Bonferroni adjustment. A threshold value for significance was set as $p < 0.05$.

PSG and DO are considered as the gold standard measure in pediatric sleep studies [15,21]. In this study, the DO served as the gold standard. Validation of actigraphy and aEEG against DO were studied by epoch-by-epoch comparisons in 30 s. Finally, concordance between actigraphy/aEEG and DO was determined.

Actigraphy, aEEG and DO data were reduced to binary form (0 = wakefulness and 1 = sleep) for analyses. Data were analyzed with cross-tabulation. Sensitivity was the proportion of correctly identified sleep epochs by actigraphy/aEEG to the epochs defined as sleep by DO. Specificity was defined as the proportion of accurately identified wake epochs by actigraphy/aEEG to the DO scored wake epochs. Agreement rate, predictive value of sleep (PVS), and predictive value of wake (PVW) were calculated as described in previous studies [16,17,19].

Agreement between actigraphy/aEEG and DO was determined by calculating Cohen's Kappa. Because sleep and wake were not equally distributed in both methods, we additionally calculated prevalence-adjusted and bias-adjusted kappa (PABAK) to provide equal weights to sleep and wake [35]. Kappas (for both Cohen's and PABAK statistics) between 0.01 and 0.20 was considered as none to slight, 0.21–0.40 as fair, 0.41–0.60 as moderate, 0.61–0.80 as substantial, and 0.81–1.00 as almost perfect agreement [36].

Bland-Altman concordance plots were used to visualize the degree of agreement between sleep parameters derived from the new method (actigraphy/aEEG) with those from the gold standard (DO) [37]. The difference between the measures for each participant was fitted to lines that represent the ideal (representing perfect agreement), plus either standard deviations or time discrepancies to show each participant's deviation from the ideal. The mean of each sleep parameter with the two techniques was represented in the x-axis and differences in each sleep parameter between the two techniques were represented in the y-axis.

Sound levels were quantified by the 10th, 30th, 50th, 70th, and 90th percentile noise levels in dB (L_{10}-L_{90}). L_{max} was the highest sound level in any of the short measuring intervals. Percentages of measurements that exceeded American Academy of Pediatrics (AAP) recommendations (>65 dB) were also measured. AAP recommends not to exceed an hourly equivalent continuous sound level (L_{eq}) of 45 dB and an hourly L_{10} of 50 dB (no more than 10% of the time at greater than 50 dB) in any bed space or patient care area [38].

3. Results

3.1. Characteristics of the infants

Seventeen preterm infants born at 28-$36^{6/7}$ weeks of GA were enrolled. Characteristics of the infants are presented in Table 1. None of the included infants were on respiratory support at the time of study inclusion.

3.2. NICU environment

Mean light exposure inside the incubator from 8:00 a.m. to 8:00 p.m. was 45.6 ± 74.9 lux (median 28.6 lux), and 14.6 ± 11.6 lux (9.6 lux) from 8:00 p.m. to 8:00 a.m. Daytime light exposure was significantly higher than nighttime light exposure ($p = 0.03$).

Table 1
Demographic variables and clinical data of the study participants.

Characteristics	Mean ± Standard deviation or n (%)
Gestational age at birth, weeks	31.8 ± 2.4
Very preterm	9 (52.9)
Moderate to late preterm	8 (47.1)
Postnatal age at the time of study, weeks	33.7 ± 1.6
Birth weight, gr	1721 ± 528
Female gender	12 (70.6)
5 min APGAR score	8 ± 1
Length of NICU stay, day	38 ± 40
NTISS score at admission	13.8 ± 7.6
NTISS score at the time of the study	3.9 ± 0.8
Feeding method	
Orogastric tube	7 (41.2)
Oral feeding	10 (58.8)

GA: gestational age, NICU: Neonatal intensive care unit, NTISS: Neonatal Therapeutic Intervention Scoring System.

Mean sound exposure inside the incubator from 8:00 a.m. to 8:00 p.m. was 55.4 ± 1.2 dB, and 52.8 ± 1.8 from 8:00 p.m. to 8:00 a.m.). Daytime sound exposure was significantly higher than nighttime ($p < 0.01$). Sound level exposure of the neonates inside the incubators is presented in Supplemental Table 1S.

3.3. Cerebral function monitoring

aEEG recordings were characterized by continuous tracing (all of the scores were 1–2), cycling was recognizable (scores≥3) in 5, and a complete mature pattern was observed in 10 neonates (only 1 neonate showing a score of 2). The lower border amplitude of the aEEG remained elevated in all. Bandwidth was narrow with scores of 3 in 12 neonates and reached its maximum level in 4. An example of the tracing analysis is shown in Supplemental Fig. S1. Results of the aEEG recording using the Burdjalov scoring system are presented in Table (Supplemental Table S2**).**

The mean Burdjalov score was higher in moderate to late preterm infants (11.3 ± 0.8) as compared to very preterm infants (9.4 ± 0.5) (p: 0.004). Gestational age was positively correlated with the Burdjalov score (r_s:0.83, $p < 0.001$).

3.4. Neurological assessments

General movement assessment revealed normal for 11 and PR for 6 patients at the time of sleep assessment. Neither CS nor chaotic GMs occurred in any of the patients. The motor repertoire was scored as age adequate in 9 patients out of 17. Videos were not available at 12 weeks corrected age in 2 patients. Three patients who had PR at 32 weeks PMA were unable to show FMs at 12 weeks corrected age (Supplemental Table 3).

Bayley composite scores for all subsets were within the normal range in all (Supplemental Table S4). Bayley could not be performed in 4 patients whose families refused to come to the hospital during the COVID pandemic. There were no differences in Bayley III composite scores with respect to GA.

3.5. Sleep analysis

Actigraphy data for 24 h are given in Supplemental Table S5. Mean total sleep time detected at the medium threshold was 16.5 h in this sample of preterm infants with a mean age of 33.7 weeks.

According to the measures by actigraphy, mean activity counts per day were 664 (total activity), 322 for the night (8 p.m.-8 am), and 342 for the daytime (8 a.m.-8 pm). Ratios of day and night activity were determined as an index of expressed rhythmicity. The day/night activity ratio was 1.07, which indicates that 7% more total activity during the day than at night. All-day sleep data at a medium threshold showed no significant difference according to the feeding method of infants (Supplemental Table S6).

Sleep data were compared according to gestational maturity (very preterm vs moderate to late preterm, Supplemental Table S7). Newborns born very preterm had longer nighttime and total sleep duration compared to newborns born moderate to late preterm at all activity thresholds except for the automatic threshold. Data derived from aEEG and DO did not show significant differences according to GA of the infants (Supplemental Tables S8 and S9).

Eight-hour actigraphy data were presented at each of the low, medium, high, and automatic activity thresholds along with sleep parameters derived from DO and aEEG (Table 2).

Sleep parameters derived from actigraphy at medium threshold did not significantly differ either from DO or aEEG ($p > 0.05$). Nighttime sleep and wake durations derived from all actigraphy thresholds were not significantly different from DO and aEEG ($p > 0.05$). aEEG-derived sleep parameters showed no significant difference from DO ($p > 0.05$).

3.5.1. Epoch-by-epoch analysis

A total of n = 11252, 30-sec epochs were studied to calculate the agreement of actigraphy and aEEG against direct observation. Epoch-by-epoch agreement for all threshold settings for actigraphy and direct observation is presented in Table 3. At all thresholds, the agreement was largely equivalent with low kappas (0.14–0.17) and PABAK coefficients (0.22–0.35) for actigraphy and DO. Moderate agreement was observed between aEEG and DO according to the PABAK coefficient (0.44).

Sensitivity, specificity, PVS, and PVW of actigraphy and aEEG against direct observation were calculated at different activity thresholds (Table 4).

Sensitivity (86.4%), agreement rate (67.9%), and PVW (47.9%) for the actigraphy measurements were highest at the automatic activity threshold whereas specificity (54.5%) and predictive value of sleep (75.5%) were highest at the low activity threshold. Sensitivity (79.3%) of aEEG against DO was high whereas specificity (54.3%) was moderate.

Sensitivity, specificity, PVS, PVW, and agreement rates of actigraphy against aEEG were calculated at each activity threshold (Table 5).

Sensitivity (86.1%) and agreement rate (68%) and PVW (46.4%) for the actigraphy measurements were highest at the automatic activity threshold. Specificity (51.5%) and PVS (74.4%) were highest at the low activity threshold.

3.5.2. Concordance between sleep parameters

Bland-Altman statistics for all activity thresholds compared to DO are presented in Table 6. There was no significant bias in sleep parameters between DO and actigraphy at the medium activity threshold. The differences in sleep parameters between actigraphy at the medium threshold and DO fell between the limits of agreement for almost all participants (Fig. 2). Actigraphy at medium threshold tend to underestimate sleep and overestimate wake, though the mean differences are small (between 3 and 6 min) as shown in Table 6. Bland-Altman plots and mean differences for the comparison of aEEG and DO and actigraphy medium threshold and aEEG for sleep efficiency, sleep time, and awake time are shown in Supplemental Figs. S2–4 and Tables S10–11. AEEG and DO-derived sleep parameters showed no significant bias. Similarly, sleep parameters of actigraphy at the medium threshold were not statistically different compared to the aEEG.

3.5.3. Sleep and the NICU environment

There was no significant association of day and nighttime sound

Table 2
Sleep parameters derived from 8-h matched-data by direct behavioral observation, actigraphy for all activity thresholds and aEEG.

	Direct Observation (DO)	Actigraphy Activity Threshold				aEEG
		Low	Medium	High	Automatic	
Daytime sleep (min)	160.2 ± 28.3	195 ± 22.4*†	157.1 ± 35.9	181.7 ± 27.1*†	132.6 ± 35.7*	150 ± 33
Nighttime sleep (min)	172.1 ± 42.1	200.8 ± 23.8	168.5 ± 34.8	190.8 ± 29.2	144.4 ± 39.8	185 ± 39.3
Total sleep (min)	332.3 ± 44.7	395.8 ± 37.7*†	325.5 ± 62.5	378.5 ± 52.9	277.1 ± 68.1*	332.7 ± 38.2
Daytime wake (min)	79.7 ± 28.3	45.0 ± 22.4*†	82.9 ± 35.9	58.2 ± 27.1*†	107.3 ± 35.7*	90 ± 33
Nighttime wake (min)	67.9 ± 42.1	39.1 ± 23.8	71.4 ± 34.8	49.1 ± 29.2	95.5 ± 39.8	55 ± 39.3
Total wake (min)	147.6 ± 44.7	84.1 ± 37.7*†	154.4 ± 62.5	101.5 ± 52.9	202.9 ± 68.1*	147.2 ± 38.2
Sleep efficiency (%)	69.2 ± 9.3	82.4 ± 7.8*†	67.7 ± 13	78.8 ± 11	57.7 ± 14.1*	69.2 ± 7.9

Data were represented as mean ± standard deviation.
*significantly different from DO $p < 0.05$.
†significantly different from aEEG $p < 0.05$.

Table 3
Epoch-by-epoch agreement for all activity threshold levels for actigraphy and aEEG with direct observation.

	Actigraphy Activity Threshold				aEEG
	Low	Medium	High	Automatic	
%Agreement	61.1	64.2	66.6	67.9	72
Kappa[a]	0.170	0.162	0.144	0.155	0.345
PABAK[b]	0.220	0.283	0.331	0.357	0.440

[a] Cohen's kappa.
[b] Prevalance and Bias Adjusted Kappa.

and light levels with the sleep parameters determined by aEEG and direct observation.

Mean sound levels at daytime showed a negative correlation with total daytime sleep (r_s: −0.55, $p = 0.02$) and a positive correlation (r_s: 0.55, $p = 0.02$) with daytime wake duration measured by actigraphy. Other sleep parameters by actigraphy did not show a significant association with the sound and light levels in the NICU.

4. Discussion

Our study compared epoch-by-epoch gold standard DO scores and actigraphy/aEEG to assess the effectiveness of actigraphy for predicting sleep/wake states in preterm-born infants after 32 weeks of age in the NICU setting. The sensitivity and AR of actigraphy were highest at the automatic threshold wherein the lowest specificity and PVS were also observed. Specificity (27.3–54.5%) and PVW (40.9–47.9%) were low across all the thresholds whereas sensitivity (64.1–86.4%), PVS (72–75.5%) and AR (61.1–67.9%) was moderate-high at each of the thresholds. Comparison of aEEG against DO revealed similar results with a sensitivity of 80.1%, specificity of 54.3%, and AR of 72%, however, slightly better agreement between the two methods was observed. Actigraphy medium threshold showed no significant bias compared to DO and aEEG. Our results add to the limited knowledge about the utility of actigraphy for depicting sleep/wake states in typically developing preterm infants. In addition, aEEG may be an alternative adjunctive method to actigraphy for the evaluation of sleep/wake states of preterm infants in the NICU setting.

To our knowledge, our study is one of the few studies performed in very young infants born preterm [15–21]. We performed sleep assessments at a mean of 34 weeks of PMA. Only a few studies used actigraphy for sleep assessments at such a young age when the infants were still in the NICU. In the study of Sung and others actigraphy was compared with direct observation in preterm

Table 4
Sensitivity and specificity analysis of actigraphy and aEEG against direct observation.

	Actigraphy Activity Threshold				aEEG
	Low	Medium	High	Automatic	
Sensitivity, % (CI)	64.1 (63–65.1)	74.4 (73.5–75.4)	83.4 (82.5–84.2)	86.4 (85.7–87.2)	80.1 (79.2–81)
Specificity, %(CI)	54.5 (52.8–56.1)	41.7 (40.1–43.4)	29.8 (28.3–31.3)	27.3 (25.8–28.8)	54.3 (52.7–56)
PVS, %(CI)	75.5 (74.4–76.5)	73.6 (72.6–74.6)	72.2 (71.2–73.1)	72 (71.3–73.1)	79.3 (78.4–80.2)
PVW, %(CI)	40.9 (39.5–42.4)	42.8 (41.1–44.4)	45 (43–47.1)	47.9 (45.7–50.1)	55.6 (53.9–57.2)

CI: Confidence Interval.
PVS: Predictive value for sleep.
PVW: Predictive value for wakefulness.

Table 5
Actigraphy against aEEG.

	Actigraphy Activity Threshold			
	Low	Medium	High	Automatic
Sensitivity, % (CI)	62.5 (61.4–63.6)	73.6 (72.6–74.6)	83.4 (82.5–84.2)	86.1 (85.4–86.9)
Specificity, % (CI)	51.5 (49.8–53.1)	40.2 (38.6–41.9)	30 (28.5–31.6)	26.7 (25.3–28.3)
PVS, % (CI)	74.4 (73.3–75.5)	73.5 (72.6–74.5)	72.9 (72.0–73.8)	72.5 (71.6–73.4)
PVW, % (CI)	37.8 (36.4–39.2)	40.3 (38.7–42.0)	44.4 (42.4–46.5)	46.4 (44.2–48.6)
AR	59.1	63.3	66.9	68

CI: Confidence Interval.
PVS: Predictive value for sleep.
PVW: Predictive value for wakefulness.
AR: Agreement rate.

Table 6
Bland-Altman statistics for actigraphy compared to direct observation at all threshold settings.

		Actigraphy threshold level			
		Low	Medium	High	Automatic
Daytime sleep	Mean Difference	−34.70*	3.23	−21.47*	27.64*
	SD Difference	28.58	32.25	26.14	31.67
	Lower Limit	−90.73	−59.98	−72.71	−34.44
	Upper limit	21.31	66.45	29.77	89.73
Nighttime sleep	Mean Difference	−28.82	3.52	−18.82	27.64
	SD Difference	43.85	52.52	47.05	58.71
	Lower Limit	−114.78	−99.41	−111.04	−87.42
	Upper limit	57.13	106.47	73.40	142.72
Total sleep time	Mean Difference	−63.52*	6.76	−46.17	55.29*
	SD Difference	48.50	62.44	59.14	69.15
	Lower Limit	−158.59	−115.63	−162.10	−80.25
	Upper limit	31.53	129.16	69.75	190.83
Daytime awake	Mean Difference	34.70*	−3.23	21.47*	−27.64*
	SD Difference	28.58	32.25	26.14	31.67
	Lower Limit	−21.31	−66.45	−29.77	−89.73
	Upper limit	90.73	59.98	72.71	34.44
Nighttime awake	Mean Difference	28.82	−3.52	18.83	−27.64
	SD Difference	43.85	52.52	47.05	58.71
	Lower Limit	−57.13	−106.47	−73.40	−142.72
	Upper limit	114.78	99.41	111.04	87.42
Total awake time	Mean Difference	63.52*	−6.76	46.17	−55.29*
	SD Difference	48.50	62.44	59.14	69.15
	Lower Limit	−31.53	−129.16	−69.75	−190.83
	Upper limit	158.59	115.63	162.10	80.25
Total sleep efficiency	Mean Difference	−13.21*	1.42	−9.60	11.53*
	SD Difference	10.10	13.00	12.31	14.40
	Lower Limit	−33.01	−24.06	−33.73	−16.69
	Upper limit	6.58	26.91	14.52	39.76

*$p < 0.05$.

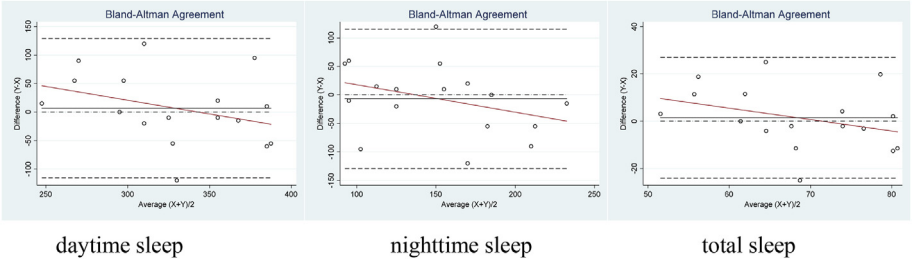

Fig. 2. Bland Altman plots for the comparison of each sleep parameter. The mean of the sleep parameters with actigraphy and direct observation are represented in the x-axis and the differences (indicating mean biases) between the two techniques are represented in the y-axis. Each dot indicates a subject. The dashed lines below and above the reference lines (the dashed line passing through zero and indicating perfect agreement) represent ± standard deviations.

infants at 2 weeks of postnatal age [17]. The highest agreement rate (84.5–88.9%), PVS (91.3–95.6%), and sensitivity (88.2–96.8%) was found in the low activity threshold. Specificity and PVW were low suggesting that actigraphy is not reliable for determining wakefulness but can be used to indicate sleep in preterm infants in the NICU settings. Another study for validation of actigraphy in preterm infants against DO demonstrated the highest sensitivity, AR, and lowest specificity in high and automatic threshold in a sample of preterm born infants [21]. The highest specificity and PVS were in the low threshold. Their results were very similar to our results indicating that actigraphy can be used to detect sleep in preterm infants in NICU settings with high sensitivity, though concerns related to the detection of wake episodes remain due to low specificity.

Studies investigated the utility of actigraphy in the assessment of preterm infants' sleep patterns since gold standard methods such as polysomnography or direct behavioral observation are time-consuming and complex but still there is a need for objective sleep assessment of preterm infants in the NICU [19,20]. Validity studies revealed conflicting results. The highest values for metrics at different thresholds differ in each study which might be attributed to the study methods and study population (term, late preterm or moderate preterm, very preterm, healthy vs sick) [16–19]. The diverse methodology of the existing studies renders the results non-comparable but in general, even though studies demonstrate high sensitivity of the actigraphy with the gold standard method, specificity remains low, which might limit the validity of actigraphy in preterm infants [39,40].

It has been suggested that robust validation of actigraphy cannot be achieved without adequate representation of sleep and wake epochs [19]. Low specificity values might be expected because of the unequal distribution of wake and sleep epochs during recording time [16,21]. However, a balanced representation of wake and sleep may not be possible; particularly in preterm infants, because

infants spent most of their time in sleep [41]. In line with his, available studies recorded limited wake periods [16–18,21]. Longer recording time is suggested to capture more wake data and our study reported a substantial amount of recorded epoch number in preterm infants. In contrast to many other studies including interfeeding periods exclusively, which may cause over-representation of sleep data, our study included feeding periods, as well [17,19]. Nevertheless, equal distribution of wake and sleep epochs was not achieved as reported in other studies. To support the validity of actigraphy, it has been suggested that other reliability statistics such as Kappa estimates of reliability, correlation, and Bland Altman plots should be reported [40]. According to the Kappa estimates, in our study, there was a low agreement between actigraphy vs DO, and aEEG vs DO. However, after adjusting for the prevalence of sleep and wake data, the PABAK coefficient showed a moderate agreement between aEEG and actigraphy. In addition, Bland Altman plots revealed that actigraphy compared to DO underestimated sleep-related parameters while overestimating wake-related parameters, which means that actigraphy failed to detect some sleep episodes, while activity is increased or differentiate wake while activity is decreased. Yet, the mean differences in sleep parameters were not significant at the medium threshold, indicating that actigraphy at the medium threshold can be used to determine the sleep pattern of preterm infants in the NICU. Interestingly, actigraphy-derived sleep parameters showed also no significant bias compared to the aEEG. These results may support the use of aEEG as an adjunct to actigraphy at the medium threshold in estimating sleep and wake cycles of preterm infants in the NICU setting.

Even though questions related to the reliability of actigraphy in detecting sleep cycles of neonates exist, actigraphy has been proposed to be useful in detecting irritability of the newborns in the NICU [19]. In line with this, we detected that with increasing daytime sound levels, the activity of neonates detected as wake by actigraphy was increased which was not observed in other methods. Nighttime sound levels did not seem to increase the irritability of the infants, most probably because of the lower levels compared to the daytime. It is suggested to incorporate non-obtrusive sleep measurements in daily ward round assessments in order to underline the importance of sleep in neonatal health [42]. Considering actigraphy as an objective measure of infant irritability may help in optimizing the NICU environment to promote sleep.

Since actigraphy is a non-intrusive method for sleep estimation, efforts to increase reliability in infants are still ongoing. New methodologies such as applying adjustments of sleep estimation to reduce disagreement between the algorithms may increase the reliability of actigraphy for infants [43]. A newly developed scoring algorithm recently presented increased specificity of actigraphy in adults [44]. The value of changing algorithms in the actigraphic assessment of sleep in preterm infants has to be studied to increase specificity in preterm infants.

Our study is unique in investigating the validity of actigraphy against both DO and aEEG performed in typically developing preterm infants in the NICU setting. However, the study has some limitations. It is obvious that actigraphy cannot be used interchangeably as a gold standard method. Considering the low specificity, actigraphy might not be an appropriate diagnostic tool for preterm infants. Even in healthy adults, the specificity of actigraphy may remain low [45]. This is mainly due to the reason that actigraphy tends to misclassify motionless wake epochs as sleep. However, actigraphy might still be useful for monitoring sleep patterns in this highly vulnerable group of infants in NICU care, because further agreement analysis revealed no significant bias in sleep parameters at the medium threshold. Limitations of actigraphy in the assessment of sleep patterns of preterm infants have been defined. Since actigraphy is basically a motion sensor, it may not discriminate between periods of quiet wakefulness and sleep or active sleep and wakefulness. High number of body movements has been noted in preterm infants, especially during active sleep. Besides, actigraphy is prone to artifacts such as external motion. Despite the limitations, actigraphy is still being investigated to document the sleep patterns of preterm infants [20]. To overcome these challenges, documentation of sleep and wake periods is necessary by complementary methods such as sleep diaries. In our study, aEEG showed considerable sensitivity and AR, therefore, aEEG may be an alternative complementary method to support actigraphy findings in the NICU settings.

5. Conclusion

In conclusion, actigraphy is a noninvasive practical tool for sleep assessments of preterm infants in the NICU with high sensitivity and AR and nonsignificant bias at the medium threshold. However, results should be interpreted cautiously since specificity and PVW are low. Adjunctive methods such as sleep diaries or aEEG, which is another bedside tool, might add to the clinical application of actigraphy in the NICU.

Funding

This work was supported by Marmara University Scientific Research Commission BAPKO (SAG-C-TUP-131217-0647).

CRediT authorship contribution statement

Özge Ülgen: study conception and design, data collection, Writing – original draft, commented on final version, All authors read and approved the final version of the manuscript. **Hatice Ezgi Barış:** Formal analysis, Writing – original draft, commented on final version, All authors read and approved the final version of the manuscript. **Öykü Özbörü Aşkan:** study conception and design, data collection, commented on final version. **Selda Küçük Akdere:** data collection, commented on final version, All authors read and approved the final version of the manuscript. **Can Ilgın:** Formal analysis, commented on final version, All authors read and approved the final version of the manuscript. **Hülya Özdemir:** study conception and design, data collection, commented on final version, All authors read and approved the final version of the manuscript. **Nural Bekiroğlu:** Formal analysis, commented on final version, All authors read and approved the final version of the manuscript. **Kıvılcım Gücüyener:** study conception and design, data collection, Formal analysis, commented on final version, All authors read and approved the final version of the manuscript. **Eren Özek:** study conception and design, data collection, commented on final version, All authors read and approved the final version of the manuscript. **Perran Boran:** study conception and design, Formal analysis, Writing – original draft, commented on final version, All authors read and approved the final version of the manuscript.

Acknowledgments

We are grateful to all patients who participated in this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.sleep.2022.11.020.

References

[1] Chawanpaiboon S, Vogel JP, Moller AB, et al. Global, regional, and national estimates of levels of preterm birth in 2014: a systematic review and modelling analysis. Lancet Global Health 2019;71:e37–46. https://doi.org/10.1016/S2214-109X(18)30451-0.

[2] Yalçin SS, Boran P, Tezel B, et al. Effects of the COVID-19 pandemic on perinatal outcomes: a retrospective cohort study from Turkey. BMC Pregnancy Childbirth 2022;22:51. https://doi.org/10.1186/s12884-021-04349-5.

[3] Bell EF, Hintz SR, Hansen NI, et al. Mortality, in-hospital morbidity, care practices, and 2-year outcomes for extremely preterm infants in the US, 2013-2018. JAMA 2022;327:248–63. https://doi.org/10.1001/jama.2021.23580.

[4] Trickett J, Hill C, Austin T, Johnson S. The impact of preterm birth on sleep through infancy, childhood and adolescence and its implications. Children 2022;9:626. https://doi.org/10.3390/children9050626.

[5] Bennet L, Walker DW, Horne RSC. Waking up too early - the consequences of preterm birth on sleep development. J Physiol 2018;596:5687–708.

[6] Kurth S, Dean 3rd DC, Achermann P, et al. Increased sleep depth in developing neural networks: new insights from sleep restriction in children. Front Hum Neurosci 2016;10:456. https://doi.org/10.3389/fnhum.2016.00456.

[7] Levy J, Hassan F, Plegue MA, et al. Impact of hands-on care on infant sleep in the neonatal intensive care unit. Pediatr Pulmonol 2017;52:84–90. https://doi.org/10.1002/ppul.23513.

[8] van den Hoogen A, Teunis CJ, et al. How to improve sleep in a neonatal intensive care unit: a systematic review. Early Hum Dev 2017;113:78–86. https://doi.org/10.1016/j.earlhumdev.2017.07.002.

[9] Gogou M, Haidopoulou K, Pavlou E. Sleep and prematurity: sleep outcomes in preterm children and influencing factors. World J Pediatr 2019;15:209–18. https://doi.org/10.1007/s12519-019-00240-8.

[10] Stangenes KM, Fevang SK, Grundt J, et al. Children born extremely preterm had different sleeping habits at 11 years of age and more childhood sleep problems than term-born children. Acta Paediatr 2017;106:1966–72. https://doi.org/10.1111/apa.13991.

[11] Luijk MPCM, Kocevska D, Tham EKH, et al. Gestational age at birth and sleep duration in early childhood in three population-based cohorts. Sleep Med X 2019;1:100002. https://doi.org/10.1016/j.sleepx.2019.

[12] Lyu J, Groeger JA, Barnett AL, et al. Associations between gestational age and childhood sleep: a national retrospective cohort study. BMC Med 2022;20:253. https://doi.org/10.1186/s12916-022-02443-9.

[13] Bulut O, Cam S, Ovali F. Impact of sleep behaviors on social and emotional problems in three-year-old children born prematurely. Sleep Med 2020;74:173–8. https://doi.org/10.1016/j.sleep.2020.08.004.

[14] Werth J, Atallah L, Andriessen P, et al. Unobtrusive sleep state measurements in preterm infants - a review. Sleep Med Rev 2017;32:109–22. https://doi.org/10.1016/j.smrv.2016.03.005.

[15] Meltzer LJ, Montgomery-Downs HE, Insana SP, Walsh CM. Use of actigraphy for assessment in pediatric sleep research. Sleep Med Rev 2012;16:463–75. https://doi.org/10.1016/j.smrv.2011.10.002.

[16] So K, Buckley P, Adamson TM, Horne RS. Actigraphy correctly predicts sleep behavior in infants who are younger than six months, when compared with polysomnography. Pediatr Res 2005;58(4):761–5. https://doi.org/10.1203/01.PDR.0000180568.97221.56.

[17] Sung M, Adamson MT, Horne RSC. Validation of actigraphy for determining sleep and wake in preterm infants. Acta Pediatrica 2009;98:52–7. https://doi.org/10.1111/j.1651-2227.2008.01002.x.

[18] Rioualen S, Roué JM, Lefranc J, et al. Actigraphy is not a reliable method for measuring sleep patterns in neonates. Acta Paediatr 2015;104:e478–82. https://doi.org/10.1111/apa.13088.

[19] Derbin M, McKenna L, Chin D, et al. Actigraphy: metrics reveal it is not a valid tool for determining sleep in neonates. J Sleep Res 2022;31:e13444. https://doi.org/10.1111/jsr.13444.

[20] Unno M, Morisaki T, Kinoshita M, et al. Validation of actigraphy in hospitalised newborn infants using video polysomnography. J Sleep Res 2022;31:e13437. https://doi.org/10.1111/jsr.13437.

[21] Yang S, Yang A, Chang Y. Validation of actiwatch for assessment of sleep-wake states in preterm infants. Asian Nurs Res 2014;3:201–6. https://doi.org/10.1016/j.anr.2014.06.002.

[22] Deshpande P, McNamara PJ, Hahn C, et al. A practical approach toward interpretation of amplitude integrated electroencephalography in preterm infants. Eur J Pediatr 2022;181:2187–200. https://doi.org/10.1007/s00431-022-04428-5.

[23] Li XF, Zhou YX, Zhang L. Newborns' sleep-wake cycle development on amplitude integrated electroencephalography. World J Pediatr 2016;12:327–34. https://doi.org/10.1007/s12519-016-0026-x.

[24] Engle WA. The American Academy of pediatrics committee on fetus and newborn. Age terminology during the perinatal period. Pediatrics 2004;114:1362–4. https://doi.org/10.1542/peds.2004-1915.

[25] March of Dimes PMNCH. In: Howson CP, Kinney MV, Lawn JE, editors. Save the children, WHO. Born too soon: the global action report on preterm birth. Geneva: World Health Organization; 2012.

[26] Gray JE, Richardson DK, McCormick MC, et al. Neonatal therapeutic intervention scoring system: a therapy-based severity-of-illness index. Pediatrics 1992;90:561–7.

[27] Grigg-Damberger MM. The visual scoring of sleep in infants 0 to 2 Months of age. J Clin Sleep Med 2016;12:429–45. https://doi.org/10.5664/jcsm.5600.

[28] Figueiredo B, Dias CC, Pinto TM, Field T. Infant sleep-wake behaviors at two weeks, three and six months. Infant Behav Dev 2016;44:169–78. https://doi.org/10.1016/j.infbeh.2016.06.011.

[29] Burdjalov VF, Baumgart S, Spitzer AR. Cerebral function monitoring: a new scoring system for the evaluation of brain maturation in neonates. Pediatrics 2003;112:855–61. https://doi.org/10.1542/peds.112.4.855.

[30] Parmalee AH, Stern E. Development of states in infants. In: Clemente CD, Purpura DP, Mayer FE, editors. Sleep and the maturing nervous system. New York, NY: Academic Press; 1972. p. 199–228.

[31] Gill NE, Behnke M, Conlon M, et al. Effect of nonnutritive sucking on behavioral state in preterm infants before feeding. Nurs Res 1988;37:347–50.

[32] Einspieler C, Prechtl H, Bos A, Ferrari F, Cioni G. Prechtl's method on the qualitative assessment of general movements inpreterm, term and young infants. London: MacKeith Press, distributed by Cambridge University Press; 2004.

[33] Einspieler C, Marschik PB, Pansy J, et al. The general movement optimality score: a detailed assessment of general movements during preterm and term age. Dev Med Child Neurol 2016;58:361–8. https://doi.org/10.1111/dmcn.12923.

[34] Jary S, Whitelaw A, Walløe L, Thoresen M. Comparison of Bayley-2 and Bayley-3 scores at 18 months in term infants following neonatal encephalopathy and therapeutic hypothermia. Dev Med Child Neurol 2013;55:1053–9. https://doi.org/10.1111/dmcn.12208.

[35] Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. J Clin Epidemiol 1993;46:423–9. https://doi.org/10.1016/0895-4356(93)90018-v.

[36] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.

[37] Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res 1999;8:135–60. https://doi.org/10.1177/096228029900800204.

[38] Noise: a hazard for the fetus and newborn. American Academy of pediatrics (AAP). Committee on environmental health. Pediatrics 1997;100(4):724–7.

[39] Meltzer LJ, Montgomery-Downs HE, Insana SP, Walsh CM. Use of actigraphy for assessment in pediatric sleep research. Sleep Med Rev 2012;16:463–75. https://doi.org/10.1016/j.smrv.2011.10.002.

[40] Sadeh A. The role and validity of actigraphy in sleep medicine: an update. Sleep Med Rev 2011;15:259–67. https://doi.org/10.1016/j.smrv.2010.10.001.

[41] Ardura J, Andrés J, Aldana J, Revilla MA. Development of sleep-wakefulness rhythm in premature babies. Acta Paediatr 1995;84:484–9. https://doi.org/10.1111/j.1651-2227.1995.tb13679.x.

[42] van den Hoogen A, Teunis CJ, Shellhaas RA, et al. How to improve sleep in a neonatal intensive care unit: a systematic review. Early Hum Dev 2017;113:78–86. https://doi.org/10.1016/j.earlhumdev.2017.07.002.

[43] Schoch SF, Jenni OG, Kohler M, Kurth S. Actimetry in infant sleep research: an approach to facilitate comparability. Sleep 2019;42:zsz083. https://doi.org/10.1093/sleep/zsz083.

[44] Lüdtke S, Hermann W, Kirste T, et al. An algorithm for actigraphy-based sleep/wake scoring: comparison with polysomnography. Clin Neurophysiol 2021;132:137–45. https://doi.org/10.1016/j.clinph.2020.10.019.

[45] Marino M, Li Y, Rueschman MN, et al. Measuring sleep: accuracy, sensitivity, and specificity of wrist actigraphy compared to polysomnography. Sleep 2013;36:1747–55. https://doi.org/10.5665/sleep.3142.