ORIGINAL ARTICLE

Neonatal Sleep–Wake Analyses Predict 18-month Neurodevelopmental Outcomes

Renée A. Shellhaas, MD, MS[1,3]; Joseph W. Burns, PhD[2]; Fauziya Hassan, MD, MS[1,3]; Martha D. Carlson, MD, PhD[1]; John D.E. Barks, MD[1]; Ronald D. Chervin, MD, MS[3,4]

[1]Department of Pediatrics and Communicable Diseases, University of Michigan, Ann Arbor, MI; [2]Michigan Tech Research Institute, Michigan Technological University, Ann Arbor, MI; [3]Sleep Disorders Center, University of Michigan, Ann Arbor, MI; [4]Department of Neurology, University of Michigan, Ann Arbor, MI

Objectives: The neurological examination of critically ill neonates is largely limited to reflexive behavior. The exam often ignores sleep–wake physiology that may reflect brain integrity and influence long-term outcomes. We assessed whether polysomnography and concurrent cerebral near-infrared spectroscopy (NIRS) might improve prediction of 18-month neurodevelopmental outcomes.

Methods: Term newborns with suspected seizures underwent standardized neurologic examinations to generate Thompson scores and had 12-hour bedside polysomnography with concurrent cerebral NIRS. For each infant, the distribution of sleep–wake stages and electroencephalogram delta power were computed. NIRS-derived fractional tissue oxygen extraction (FTOE) was calculated across sleep–wake stages. At age 18–22 months, surviving participants were evaluated with Bayley Scales of Infant Development (Bayley-III), 3rd edition.

Results: Twenty-nine participants completed Bayley-III. Increased newborn time in quiet sleep predicted worse 18-month cognitive and motor scores (robust regression models, adjusted $r^2 = 0.22$, $p = .007$, and 0.27, .004, respectively). Decreased 0.5–2 Hz electroencephalograph (EEG) power during quiet sleep predicted worse 18-month language and motor scores (adjusted $r^2 = 0.25$, $p = .0005$, and 0.33, .001, respectively). Predictive values remained significant after adjustment for neonatal Thompson scores or exposure to phenobarbital. Similarly, an attenuated difference in FTOE, between neonatal wakefulness and quiet sleep, predicted worse 18-month cognitive, language, and motor scores in adjusted analyses (each $p < .05$).

Conclusions: These prospective, longitudinal data suggest that inefficient neonatal sleep—as quantified by increased time in quiet sleep, lower electroencephalogram delta power during that stage, and muted differences in FTOE between quiet sleep and wakefulness—may improve prediction of adverse long-term outcomes for newborns with neurological dysfunction.

Keywords: Neonatal polysomnography, neurodevelopmental outcomes, near-infrared spectroscopy, neonatal intensive care.

Statement of significance

In this prospective study of bedside polysomnography and near-infrared spectroscopy for newborns at risk for seizures, inefficient neonatal sleep patterns (increased proportion of quiet sleep, lower EEG power during that stage, and muted differences in brain oxygen metabolism between quiet sleep and wakefulness) were independent predictors of 18-month neurodevelopmental outcome. Neonatal sleep is a novel, independent marker of brain function and can enhance prediction of risk for neurodevelopment disability. Importantly, several innovative measures of neonatal sleep neurophysiology remained clear predictors of 18-month outcomes even after adjustment for neonatal neurological examination scores and for exposure to phenobarbital.

INTRODUCTION

Objective measures of neurologic function for newborn infants are lacking in clinical practice and research protocols. Although sleep is a highly sophisticated brain function and is known to be critical for learning and development,[1–3] it is rarely included in the newborn clinical neurological assessment. Yet, sleep can be monitored directly and objectively with polysomnography, and patterns of brain oxygen metabolism during sleep–wake stages may be reflected by cerebral near-infrared spectroscopy (NIRS).

Our initial observations from a cross-sectional study suggested that innovative, objective analyses derived from polysomnograms can provide a window on concurrent neonatal brain function for newborns who require neonatal intensive care. Specifically, increased time in quiet sleep and decreased electroencephalograph (EEG) delta-frequency power, along with lower sleep–wake entropy, were all associated with worse neonatal neurological examination scores.[4] These findings suggested that newborns with abnormal neurological function could have amplified "pressure" to generate quiet sleep. Evidence appeared to include the higher proportion of quiet sleep, perhaps because slow wave (low delta) EEG activity was less intense in this state, decreased sleep–wake state entropy, and propensity to terminate active sleep with quiet sleep rather than wakefulness or indeterminate sleep.

Neonatal sleep measures can distinguish healthy newborns from those with hypoxic-ischemic encephalopathy (HIE),[5] and EEG delta power differs between newborns born small for gestational age and those with normal birth weights.[6] Those previous observations also support the hypothesis that newborn sleep may reflect brain functional integrity. Based on behavioral observations, others have reported that, for preterm infants, organization of neonatal sleep–wake cycling is associated with executive functioning and verbal IQ scores at age 5 years.[7] For older infants and children, abnormal sleep has neurocognitive consequences. Among otherwise healthy infants, parental report of frequent snoring—a sign of abnormal sleep—has been associated with measurably lower cognitive scores compared with infants who do not snore.[8] In a population-based study, parent-reported symptoms of sleep disordered breathing, even when they occurred transiently during the first year of life and then resolved, were associated with behavior challenges at ages 4- and 7-years.[9] Most of these published data relied on subjective reporting of sleep-related behaviors as predictors of neurobehavioral outcomes. We hypothesized that advanced polysomnographic analyses, beyond standard measures of

sleep-disordered breathing, can provide objective sleep measures that will be applicable across a range of infants at risk for neurodevelopmental disabilities.

In addition to EEG and polysomnography, neonatal neuromonitoring often includes assessment of brain oxygen metabolism through NIRS.[10] In a pilot study of concurrent polysomnography and NIRS, we reported that brain oxygen metabolism, measured by NIRS, varies across sleep–wake stages.[11] We sought to confirm this observation with a larger sample of newborns, as well as to determine whether cycling of brain oxygen metabolism across sleep–wake stages is a sign of brain health and predicts favorable long-term neurodevelopmental outcomes.

We hypothesized that among newborns who require intensive care, neonatal sleep measures can add to current standard predictors of long-term neurodevelopmental outcomes. We therefore studied measures of quiet sleep (e.g., percent time in quiet sleep and EEG delta power during that stage), variations in brain oxygen metabolism across sleep–wake stages (measured by NIRS), and a standard bedside neurological examination assessment, as possible predictors of 18-month neurodevelopmental outcomes for newborns at risk for cerebral dysfunction.

METHODS

This research was approved by our Institutional Review Board, and a parent of every participant provided written informed consent. Newborn infants (≥35 weeks gestation) who were determined clinically to be at risk for seizures, according to published guidelines,[12] were eligible for this multimodality neuromonitoring study that included conventional video EEG, cerebral NIRS, and a 12-hour full bedside polysomnogram. Preliminary analyses from some participants were presented elsewhere.[4,11,13] Infants were enrolled between 3/2010 and 6/2014. Exclusion criteria were as follows: confirmed or suspected genetic conditions that confer an independent risk for abnormal neurodevelopment (e.g., trisomy 21), congenital malformations that predispose to sleep-disordered breathing (e.g., micrognathia), and markedly abnormal background EEG patterns, such as burst suppression, that would preclude identification of sleep–wake cycling and confer a nearly uniform adverse neurodevelopmental outcome.

Demographic and clinical information was recorded from the medical record, and the Score for Neonatal Acute Physiology, Perinatal Extension, Version II (SNAPPE-II) was calculated.[14] Standardized neurological examinations were performed by a pediatric neurologist (R.A.S.) on the day of the polysomnogram and before its results were available. Thompson scores[15,16] were calculated for each infant as the clinical standard measure of current neurological status. Thompson scores range from 0 to 22 and higher scores indicate a more abnormal examination.

Neonatal NIRS sensors (Invos 5100c, Somanetics Corp., Troy, MI) were placed over bilateral parietal head regions. During each polysomnogram, the cerebral regional oxygen saturation (rSO_2) and pulse oximetry were recorded every 5 seconds on a research computer (BedMaster, ExcelMedical, Juniper, FL). Fractional tissue oxygen extraction was calculated [$FTOE = (SaO_2 - rSO_2)/SaO_2$].

Every neonate was monitored with an attended, bedside polysomnogram in the Neonatal Intensive Care Unit (NICU). Polysomnography was recorded once the infant was medically stable. Eleven patients had HIE and eight were treated with therapeutic hypothermia (three did not meet inclusion criteria for this intervention[17] due to lack of clinical encephalopathy or transfer to our NICU after 6 hours of life). In our NICU, neonates who receive therapeutic hypothermia do not receive prophylactic sedation and only those with clinical or EEG-confirmed seizures are treated with anti-seizure medications. Newborns who received therapeutic hypothermia for HIE underwent polysomnography on day of life 3 or 4 (after 72 hours of cooling). Initial analyses did not reveal major differences between neonates with HIE and patients with other diagnoses; therefore, the data were analyzed as a single sample.

In addition to a 9-channel neonatal-montage EEG, the polysomnogram included time-locked video recording, bilateral electrooculogram, chin surface EMG, chest and abdominal excursion (inductance plethysmography), nasal pressure, nasal/oral airflow (thermocouples), snoring sensor, oxygen saturation, electrocardiogram, bilateral anterior tibialis surface EMG, and transcutaneous CO_2. Sleep–wake stage scoring was based on combined information derived from behavioral observation and other recorded polysomnographic data, all according to standard neonatal scoring rules.[18] All polysomnograms were scored off-line by a single, experienced, registered polysomnographic technologist and reviewed by a board-certified sleep medicine physician. The technologist and physician were blinded to the infant's neurological exam and SNAPPE-II score.

For each participant, objective quantitative polysomnographic analyses were undertaken, as previously described.[4] Calculated variables included the proportion of each sleep–wake stage, the entropy of the sequence of sleep–wake state transitions, and power spectra from the EEG portion of the polysomnogram (C4 → M1 channel). The Walsh spectral entropy method was employed to measure the entropy of sleep–wake transitions.[19,20] High entropy values suggest decreased predictability of the sleep–wake pattern, whereas lower values imply more regularity in the pattern. To compute EEG power spectra, the periodogram power for each 30-second polysomnogram epoch was normalized by the total periodogram power averaged over all epochs. We used the Welch method for fast Fourier transform (FFT).[21] Finally, we also calculated the Spearman correlation between low-frequency (0.5–2 Hz) EEG power and FTOE.

Long-term follow-up assessments were offered to all study participants and consisted of a Bayley Scales of Infant Development (Bayley-III), 3rd edition, at age 18–22 months, as well as a clinical assessment by a pediatric neurologist (M.D.C.) who was blinded to the polysomnogram results. For participants who did not return for Bayley-III, a dichotomous outcome was determined from examination of the available medical records or through a telephone conversation with a parent. Favorable outcome was defined as survival to ≥18 months without severe disability. For all participants, adverse outcome was defined as death or any of the following: *both* Bayley-III cognitive and language scale scores <80[22], disabling cerebral palsy (gross motor function classification scale ≥3),[23] blindness,

deafness, or epilepsy. Records were independently evaluated by two pediatric neurologists and a neonatologist (R.A.S., M.D.C., and J.D.B.) to assign the dichotomous outcome and consensus was reached in all cases.

Statistical Analysis

Spearman correlations were calculated to evaluate the associations among Thompson scores, SNAPPE-II scores, and polysomnographic and NIRS data. Inspection of scatter plots (Figure 1) suggested that some outliers were influential. Three children had relatively normal Thompson scores but severely abnormal Bayley-III results (cognitive scores <60). Medical record review revealed that two of these newborns had neonatal-onset epilepsy and continued to have uncontrolled seizures through the first years of life. The other had severe HIE and developed infantile spasms by 2 months of age. One additional child had a very abnormal Thompson score (16), but a normal developmental outcome (Bayley-III cognitive score 120). This infant had postnatally diagnosed posterior urethral valves and associated heart failure, and seizures related to sinovenous thrombosis, all of which resolved after neonatal urological surgery. Given these results, we used robust regression techniques that incorporate a weighting function to de-emphasize outliers via iteratively reweighted least squares.[24–26] Non-parametric rank sum tests were used to assess associations between neurophysiologic data and dichotomous outcomes.

Based on our previous results,[4,11] we designated three variables a priori as primary explanatory variables to be evaluated: time in quiet sleep, EEG delta frequency power (especially 0.5–2 Hz), and change in FTOE across sleep–wake stages. Additional associations were assessed in secondary analyses. As this was an initial investigation in this area of research, our priority was to maintain sensitivity to potential relationships; we did not adjust for multiple comparisons and $p < .05$ was considered significant. All calculations were performed using MATLAB (MathWorks, Natick, MA).

RESULTS

Fifty newborns completed a polysomnogram and follow-up data were available for 40, among whom 31 had favorable outcomes and 9 had adverse outcomes. Twenty-nine completed the Bayley-III assessment. There were no differences between the demographic and clinical characteristics of patients who were lost to follow-up and those who completed the study protocol. Clinical and demographic information are presented in Table 1.

Associations Between Neonatal Sleep Physiology and Concurrent Neurologic Status

Higher fraction of quiet sleep was associated with higher (worse) Thompson scores (*rho* = 0.59, $p < .0001$), as was decreased time in active sleep (*rho* = -0.39, $p = .004$). Lower sleep–wake stage entropy was also associated with higher (worse) Thompson scores (*rho* = −0.39, $p = .005$). There were no associations between the polysomnographic variables and gestational age.

Initial analyses revealed that EEG power in the low-delta frequency band (0.5–2 Hz) had clear associations with neurologic status, whereas higher frequencies did not. Thus, all further analyses focused on the EEG power at 0.5–2 Hz. Neonates with younger gestational ages had lower absolute EEG power at 0.5–2 Hz frequencies during quiet sleep (*rho* = −0.29, $p = .04$) and active sleep (*rho* = −0.34, $p = .017$). Lower delta power was also associated with higher (worse) Thompson scores (*rho*= −0.52, $p = .0001$; adjusted for gestational age: *rho*= −0.46, $p = .0008$). A larger difference in epoch normalized delta power between active and quiet sleep stages was also associated with higher (worse) Thompson scores (*rho* = 0.29, $p = .04$).

The median cerebral FTOE was lowest during quiet sleep (19.3%, IQR 13.1), and highest during wakefulness (22.3%, IQR 13.5) and active sleep (22.3%, IQR 13.0; Kruskal–Wallis test, $p < .001$). For the whole group, the absolute value of the differences in FTOE across sleep stages was small, but the magnitude of these differences varied across individuals. Within every sleep–wake stage, FTOE was lower for participants with

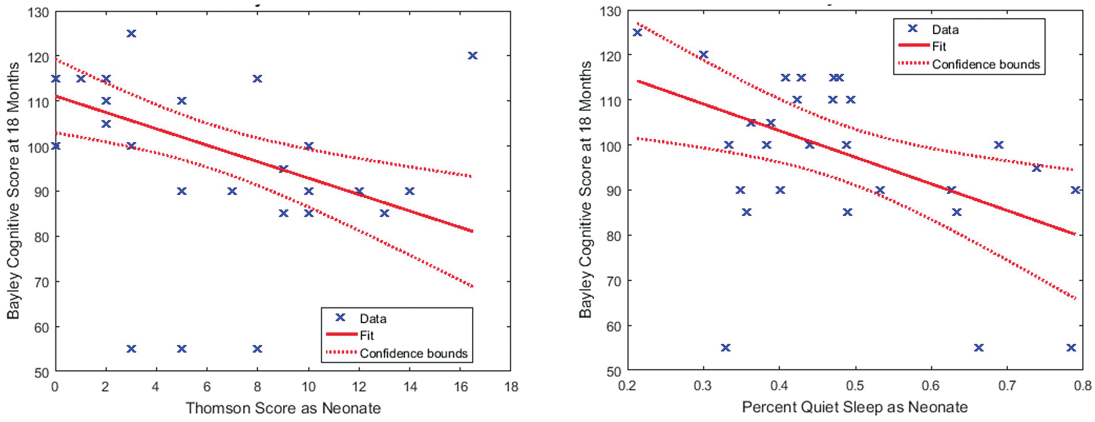

Figure 1—Neurodevelopmental outcome scores were associated with neonatal neurological examination scores and neonatal sleep parameters. (A) Lower (more normal) neonatal Thompson scores were associated with higher 18-month Bayley Scales of Infant Development, 3rd edition, cognitive scores; however, there were three outliers who received cognitive scores of 55. (B) The percent time spent in neonatal quiet sleep was associated with Bayley Scales of Infant Development, 3rd edition, cognitive scale scores; however, there were three outliers who received cognitive scores of 55, with a range of percent quiet sleep.

Table 1—Clinical and Demographic Profile of 50 Newborns who Underwent Polysomnography.

Variable	*N* = 50 with PSG	*N* = 40 with any outcome data	*N* = 29 with Bayley-III
Sex	29 males	24 males	17 males
Gestational age	39.2 ± 1.7 weeks	39.3 ± 1.5 weeks	39.6 ± 1.4 weeks
Birth weight	3361 ± 565 g	3339 ± 575 g	3419 ± 531 g
Head circumference	34.5 ± 1.9 cm	34.7 ± 2.0 cm	34.7 ± 1.7 cm
5-minute Apgar score	Median 8 [IQR 5, 10]	Median 8 [IQR 4, 10]	Median 8.5 [IQR 4.75, 10]
Thompson score	Median 4 [IQR 2, 9]	Median 4.5 [IQR 2, 10]	Median 4 [IQR 2, 9]
SNAPPE-II score	Median 24 [IQR 15, 45.5]	Median 26 [IQR 19, 49]	Median 26 [19, 49]
Primary neurologic diagnosis	HIE *N* = 15 Seizures (without obvious cause) *N* = 5 Epilepsy *N* = 5 (three epileptic encephalopathies, two benign neonatal seizures) Infection *N* = 5 Apnea *N* = 7 Arterial ischemic stroke *N* = 3 Other *N* = 9	HIE *N* = 14 Seizures (without obvious cause) *N* = 5 Epilepsy *N* = 3 (two epileptic encephalopathies; 1 benign neonatal seizure) Infection *N* = 3 Apnea *N* = 7 Stroke *N* = 2 Other *N* = 4	HIE *N* = 11 Seizures (without obvious cause) *N* = 5 Epilepsy *N* = 2 (one epileptic encephalopathy, 1 benign neonatal seizure) Apnea *N* = 4 Other *N* = 4
Therapeutic hypothermia	11	10	8
Received phenobarbital[a]	24	22	17
Bayley-III cognitive score (*N* = 28)			97 ± 19
BAYLEY-III language score (*N* = 27)			88 ± 21
Bayley-III motor score (*N* = 26)			90 ± 18

PSG = polysomnogram; HIE = hypoxic ischemic encephalopathy; Bayley-III = Bayley Scales of Infant Development, 3rd edition.
[a]Four patients received levetiracetam in addition to phenobarbital; of these, two also received fosphenytoin and one was treated with a lidocaine infusion.

younger gestational ages (FTOE in quiet sleep *rho* = −0.50, p = .001; active sleep *rho* = −0.47, p = .003; wakefulness *rho* = −0.47, p = .004). Absolute values of rSO_2 or FTOE were not associated with Thompson scores ($p > .5$ for every comparison). However, a smaller difference in FTOE between active and quiet sleep was associated with higher (worse) Thompson scores (*rho* = −0.36, p = .03).

Associations Between Neonatal Sleep Physiology and 18- to 22-month Outcomes

Within the limited range of the inclusion criteria (>35 weeks gestation), gestational age was not predictive of Bayley-III scores. Higher (more abnormal) neonatal Thompson scores predicted lower 18-month Bayley-III cognitive scores (model p = .001), but not motor (p = .03) or language (p = .16) scores (Table 2). Higher (more abnormal) SNAPPE-II scores did not predict Bayley-III scores ($p > .3$ for all comparisons). As temperature is part of the SNAPPE-II score, and most of the infants with HIE received therapeutic hypothermia, we re-evaluated SNAPPE-II scores after excluding points assigned for abnormal temperature. This adjusted SNAPPE-II score also did not predict Bayley-III results ($p > .4$ for all comparisons).

Using non-parametric tests, neither neonatal polysomnographic data nor neonatal NIRS measures were predictive of the dichotomous favorable vs. adverse outcome at age 18–22 months.

Univariate Analyses of Polysomnographic Variables

Increased time spent in neonatal quiet sleep predicted lower 18-month cognitive (p = .007), language (p = .03), and motor (p = .004) scores. Higher entropy of sleep–wake transitions was predictive of lower motor outcomes (p = .05). Diminished low frequency EEG power (0.5–2 Hz) during quiet sleep was strongly predictive of lower language (p = .0005) and motor (p = .001) scores (Table 2; Figure 2).

Adjusted Analyses of Polysomnographic Variables

As Thompson scores predicted Bayley-III scores at 18-month follow-up, we performed a series of bivariate analyses to determine whether polysomnographic variables predicted Bayley-III scores independently of Thompson scores. After adjusting for Thompson scores, increased percent time in neonatal quiet sleep remained predictive of lower cognitive (p = .03) and motor (p = .016) scale scores. Similarly, lower 0.5–2 Hz EEG power during neonatal quiet sleep remained predictive of worse 18-month language (p = .002) and motor (p = .003) scores. Lower correlation between low-frequency EEG power and FTOE also independently predicted worse motor outcomes (p = .008). In contrast, Thompson score was no longer

Table 2—Univariate Robust Regression Models Showed That Neonatal Polysomnography Variables and Neurological Examination (Thompson) Scores Predicted 18- to 22-Month Developmental Outcomes, Whereas Gestational age did not.

	Bayley-III subscale	Beta	Coefficient *p*-value	R^2	*N*
Percent quiet sleep	Cognitive	−59.12	.007	0.25	28
Percent quiet sleep	Language	−65.69	.032	0.17	27
Percent quiet sleep	Motor	−77.15	.004	0.30	26
Sleep wake Walsh spectral entropy	Cognitive	−0.70	.98	0.018	28
Sleep wake Walsh spectral entropy	Language	−54.15	.17	0.074	27
Sleep wake Walsh spectral entropy	Motor	−58.91	.050	0.24	26
Quiet sleep 0.5–2 Hz EEG power	Cognitive	23.38	.22	0.060	28
Quiet sleep 0.5–2 Hz EEG power	Language	77.85	.0005	0.39	27
Quiet sleep 0.5–2 Hz EEG power	Motor	70.35	.0012	0.36	26
Active sleep 0.5–2 Hz EEG power	Cognitive	28.8	.35	0.042	27
Active sleep 0.5–2 Hz EEG power	Language	66.6	.087	0.12	26
Active sleep 0.5–2 Hz EEG power	Motor	94.6	.010	0.26	25
Gestational age	Cognitive	1.08	.65	0.023	28
Gestational age	Language	1.47	.68	0.008	27
Gestational age	Motor	−1.87	.57	0.056	26
Thompson score	Cognitive	−1.81	.001	0.37	28
Thompson score	Language	−1.39	.16	0.08	27
Thompson score	Motor	−1.26	.10	0.18	26

Bayley-III = Bayley Scales of Infant Development, 3rd edition.

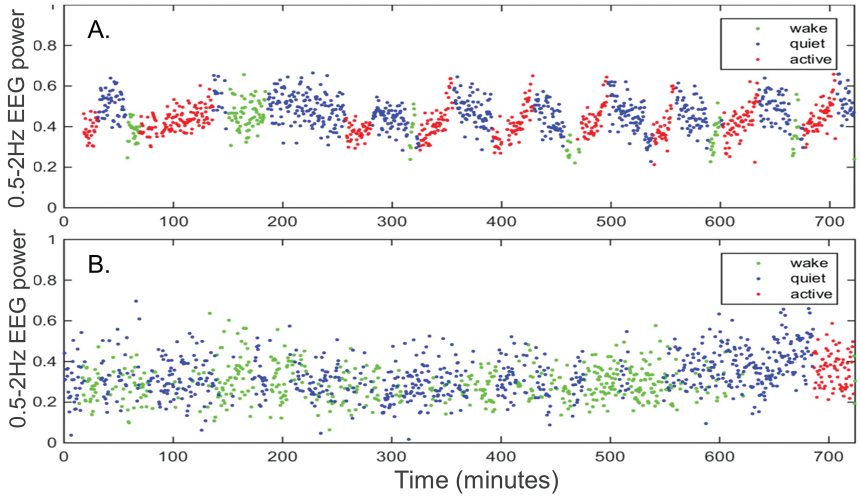

Figure 2—Neonatal EEG delta power was associated with neurodevelopmental outcome. EEG power (0.5–2 Hz) varied across sleep–wake stages for a full term newborn with HIE who had normal Bayley Scales of Infant Development, 3rd edition, scores across all domains at 18 months (A). The proportion of quiet sleep was higher and the overall delta power was lower for another term newborn with HIE who had Bayley Scales of Infant Development, 3rd edition, scores <60 in all domains at 18 months (B).

an independent predictor of outcomes after adjustment for the sleep variables.

We also explored the sensitivity of the univariate findings to neonatal phenobarbital exposure. Phenobarbital could have an effect on neonatal EEG or sleep–wake cycling and serves as a proxy for neonatal seizure diagnosis. After adjustment for phenobarbital exposure, increased time in quiet sleep remained independently predictive of worse Bayley-III cognitive ($p = .02$)

and motor (p = .01) scores; lower 0.5–2 Hz delta power in quiet sleep independently predicted worse language (p = .005) and motor (p = .002) scores; and lower entropy independently predicted worse language (p = .04) and motor (p = .02) scores. After adjustment for these sleep variables, exposure to phenobarbital was not an independent predictor of any of the outcome measures except the correlation between EEG delta power and FTOE.

Univariate Analyses of Cerebral NIRS Variables

The 18-month Bayley-III cognitive, language, and motor scale scores were not predicted by mean neonatal cerebral FTOE nor the overall FTOE variability (reflected by FTOE standard deviation) during the polysomnograms (regression model $p > .15$ for all variables). Similarly, no association emerged between the absolute FTOE value during specific sleep–wake stages and Bayley-III scores (regression model $p > .15$ for all variables). However, a reduced change in FTOE between wakefulness and quiet sleep, as a proportion of FTOE in quiet sleep [(wake FTOE − quiet sleep FTOE)/quiet sleep FTOE], was associated with lower Bayley-III language and motor scores (Table 3; Figure 3).

Adjusted Analyses of Cerebral NIRS Variables

In bivariate analyses, adjusted for Thompson scores, the relative change in FTOE between wakefulness and quiet sleep was independently predictive of 18-month cognitive (p = .033), language (p = .016), and motor (p = .003) scores. Relative change in FTOE across wakefulness and quiet sleep was also an independent predictor of cognitive (p = .028), language (p = .022), and motor (0.005) scores, after adjusting for phenobarbital exposure. In these adjusted models, Thompson score remained predictive of cognitive scores, but not language or motor scores, whereas phenobarbital exposure remained predictive of cognitive and language scores, but not motor scores.

DISCUSSION

Among late-preterm and term newborns at risk for cerebral dysfunction, we demonstrate that objective measures of neonatal brain function, recorded at the bedside through polysomnography and NIRS, have the potential to improve prediction of 18–22-month neurodevelopmental outcomes. Taken together, our results suggest that inefficient quiet sleep—more time in quiet sleep, lower EEG delta power during that state, and more attenuated changes in brain oxygen metabolism between quiet sleep and wakefulness—reflects a newborn's current neurologic status in the NICU and predicts lower 18-month neurodevelopmental outcome scores. Importantly, these novel measures of brain function remained clear predictors even after adjustment for the neonatal neurological examination scores and for exposure to phenobarbital.

EEG delta power has been reported to be higher in healthy term *versus* preterm neonates[27] and can distinguish small for gestational age (SGA) from appropriate for gestational age (AGA) term infants without cerebral dysfunction.[6] We add low-frequency EEG activity, in the 0.5–2 Hz portion of the delta spectrum, and particularly during quiet sleep, as a reflection of current neurologic status and a predictor of later motor development. An increased fraction of quiet sleep was identified, prior to the therapeutic hypothermia era, in neonates with HIE compared with healthy controls.[5] In addition, a higher quiet sleep fraction in healthy term neonates was associated with lower motor development scores at 6 months of age.[28] Our data suggest that an increased proportion of quiet sleep reflects neonatal cerebral dysfunction and independently predicts impaired neurodevelopment across all measured domains (including Bayley-III cognitive, language, and motor subscale scores).

We previously reported that cerebral NIRS data were not predictive of 18-month outcomes for a cohort of newborns treated with therapeutic hypothermia for HIE.[29] Our present results reflect that absolute values of FTOE are associated with gestational age (even across the fairly restricted range from 35 to 41 weeks gestation) and are highest during wakefulness and active sleep, but do not predict long-term outcomes. By contrast, changes in FTOE across sleep–wake stages were associated with both current neurologic exam scores and 18-month outcomes among our participants. This suggests that FTOE fluctuations may reflect normal physiologic variability during neonatal sleep–wake cycling. Of note, the rSO_2 and FTOE remained within commonly accepted normal values throughout the polysomnograms for these stable term and near-term neonates, despite their risk for seizures. It remains possible that raw NIRS data will be more predictive of long-term outcomes in other NICU sub-populations at higher risk for cerebral hypoperfusion.

Animal models have demonstrated that sleep restriction in immature animals results in altered synaptic plasticity, neuronal maturation, and subsequent behavior (summarized in Refs. [30] and [31]). In these models, both rapid eye movement (REM; akin to active sleep) and non-REM (quiet) sleep stages appear to be essential for normal brain development.[30] Our data suggest that for newborn infants with cerebral dysfunction, abnormalities in the quantity and quality of quiet sleep during the first days of life are most predictive of neurodevelopmental outcomes.

Table 3—Univariate Associations Between Changes in Cerebral FTOE Between Wakefulness and Quiet Sleep and Neurodevelopmental Outcome Scores.

	Bayley-III scale	Beta	Coefficient *p*-value	R^2	*N*
(Wake-quiet)/quiet FTOE	Cognitive score	−29.5	.056	0.18	21
(Wake-quiet)/quiet FTOE	Language score	−42.0	.025	0.25	20
(Wake-quiet)/quiet FTOE	Motor score	−43.2	.005	0.36	20

Bayley-III = Bayley Scales of Infant Development, 3rd edition; FTOE = fractional tissue oxygen extraction.

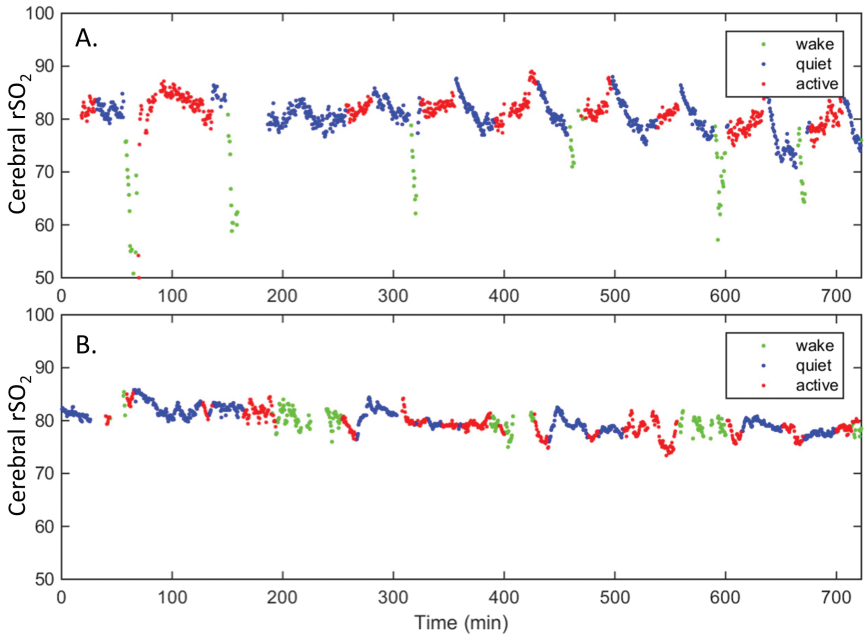

Figure 3—Changes in brain oxygen metabolism across sleep–wake stages were associated with neurodevelopmental outcome. Cerebral rSO_2 cycled across sleep–wake stages for a full term newborn with HIE who had normal Bayley Scales of Infant Development, 3rd edition, scores across all domains at 18 months (A). The changes across sleep–wake stages was muted for another term newborn with HIE who had quadriparetic cerebral palsy, no expressive language, and treatment-resistant epilepsy when she died at age 15 months (B).

Our study design allowed for intensive investigation of each newborn's sleep physiology. Thus, the sample size was, by necessity, limited. Not all patients required long-term clinical neurologic or neurodevelopmental follow-up, and the sample size for the 18-month Bayley-III was somewhat restricted. Despite the size of our cohort, our data reflect statistically significant associations between sleep data and outcomes after adjustment for important covariates. Importantly, despite the higher sample size of children for whom dichotomous outcomes were available, the measured neonatal sleep variables were only predictive of Bayley-III subscale scores and not overall outcomes. Longer-term follow-up of these children may reveal more subtle differences in executive functioning, as reported for older infants and children with symptoms of sleep-disordered breathing[8,9] or immature sleep–wake transitions in the neonatal period.[7] Such analyses should include assessments of socioeconomic status, persistent health problems, and other early-life experiences that may influence long-term neurologic function.

Our aim was clinical and highly practical—to assess whether objective neonatal sleep measures can augment existing clinical predictors of neurodevelopmental outcome for newborns at risk for developmental disability. Therefore, we did not recruit a cohort of normal control infants for this study, and our data do not permit conclusions about normal versus abnormal biology. Nonetheless, the clinical utility of our results was enhanced by the inclusion of a wide range of term and late-preterm neonates who required intensive care, in order to provide data directly relevant to this important patient group.

Previous work has presented associations between EEG and sleep findings in healthy preterm infants and neurodevelopmental outcomes. Those reports, largely based on detailed analyses of 60- to 90-minute EEG studies, suggested that brain maturation differs between children born preterm versus term, even when measured in early childhood.[32] Dysmature EEG and sleep patterns are hypothesized to reflect altered brain development in preterm neonates.[33,34] Abnormal preterm neonatal quiet sleep, as reflected by decreased EEG tracé alternant pattern during a 60-minute epoch of EEG, was previously reported to be associated with lower intelligence quotients at ages 4 months to 8 years, but this association dissipated in the setting of an attentive and enriching home environment.[35]

Bedside polysomnography is resource intensive and may not yet be available in all hospitals. Although amplitude-integrated EEG (aEEG) may reflect sleep–wake cycling,[36,37] it has never been validated against gold-standard polysomnography, and there are recognized limitations in distinguishing active sleep from wakefulness. aEEG cycling is a marker for favorable outcomes among newborns with HIE or meningitis.[36,38–40] We hypothesize that measures of sleep–wake stage entropy may parallel aEEG cycling patterns and could provide an objective method for measurement of physiologic cyclicity. To date, however, assessment of aEEG cycling has relied on subjective interpretation of the trace in order to classify absent, immature, or mature patterns. Recent data suggest that neonatal actigraphy is feasible in preterm infants (30–35 weeks gestation) at term-equivalent age and could provide objective measures of sleep–wake cycling that predict attentional difficulties later in infancy.[41] Actigraphy, while a simpler approach to monitoring sleep–wake cycling, cannot provide the detailed data required for complex signal processing of neonatal sleep states.[33] We speculate that development of automated sleep–wake cycling

measures could lead to novel, objective parameters for neonatal neurological assessment. Further detailed analyses of the patterns of cyclicity in sleep-dependent variables, such as EEG power and FTOE, may provide additional opportunities to develop objective, novel measures of neonatal brain functional integrity. Optimal analyses might combine visual and digital measures to provide a rich and detailed picture of neonatal sleep patterns.[33]

Whether inefficient quiet sleep is a reflection of an abnormal neonatal brain or may augment cerebral dysfunction in an at-risk newborn remains to be determined. It is intriguing that polysomnographic and NIRS variables across neonatal sleep–wake stages could be predictive of 18-month outcomes even after adjustment for neurological examination (Thompson) scores or phenobarbital exposure, and that neurodevelopmental outcome can be independent of illness severity (SNAPPE-II) scores. Yet, many important knowledge gaps remain. We were unable to account for potential antepartum or intrapartum contributions, including maternal, fetal, and placental factors, to neonatal seizures or encephalopathy.[42] However, we speculate that quantitative sleep data could provide much needed objective biomarkers of neonatal brain injury and recovery. Our study included only term and late-preterm infants; future work to determine the value of sleep measures for preterm infants, and particular patient populations such as neonates with growth restriction could build upon our data and those from the studies of EEG and sleep conducted prior to the era of therapeutic hypothermia for HIE.

CONCLUSIONS

Sleep of critically ill children has rarely been studied in a quantitative manner, but objective sleep measures may provide independent predictors of neurodevelopmental outcomes for newborns at risk for neurologic dysfunction.

REFERENCES

1. Ednick M, Cohen AP, McPhail GL, Beebe D, Simakajornboon N, Amin RS. A review of the effects of sleep during the first year of life on cognitive, psychomotor, and temperament development. Sleep. 2009; 32(11): 1449–1458.
2. 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(2): 93–102.
3. Picchioni D, Reith RM, Nadel JL, Smith CB. Sleep, plasticity and the pathophysiology of neurodevelopmental disorders: the potential roles of protein synthesis and other cellular processes. Brain Sci. 2014; 4(1): 150–201.
4. Shellhaas RA, Burns JW, Barks JDE, Chervin RD. Quantitative sleep stage analyses as a window to neonatal neurologic function. Neurology. 2014; 82(5):390–395.
5. Scher MS, Steppe DA, Beggarly ME, Salerno DG, Banks DL. Neonatal EEG-sleep disruption mimicking hypoxic-ischemic encephalopathy after intrapartum asphyxia. Sleep Med. 2002; 3(5): 411–415.
6. Ozdemir OM, Ergin H, Sahiner T. Electrophysiological assessment of the brain function in term SGA infants. Brain Res. 2009; 1270: 33–38.
7. Weisman O, Magori-Cohen R, Louzoun Y, Eidelman AI, Feldman R. Sleep-wake transitions in premature neonates predict early development. Pediatrics. 2011; 128(4): 706–714.
8. Piteo AM, Kennedy JD, Roberts RM, et al. Snoring and cognitive development in infancy. Sleep Med. 2011; 12(10): 981–987.
9. Bonuck K, Freeman K, Chervin RD, Xu L. Sleep-disordered breathing in a population-based cohort: behavioral outcomes at 4 and 7 years. Pediatrics. 2012; 129(4): e857–e865.
10. Sood BG, McLaughlin K, Cortez J. Near-infrared spectroscopy: applications in neonates. Semin Fetal Neonatal Med. 2015; 20(3): 164–172.
11. Shellhaas RA, Burns JW, Wiggins SA, Christensen MK, Barks JD, Chervin RD. Sleep-wake cycling and cerebral oxygen metabolism among critically ill neonates. J Child Neurol. 2014; 29(4): 530–533.
12. Shellhaas RA, Chang T, Tsuchida T, et al. The American Clinical Neurophysiology Society's guideline on continuous electroencephalography monitoring in neonates. J Clin Neurophysiol. 2011; 28(6): 611–617.
13. Sokoloff MD, Plegue MA, Chervin RD, Barks JD, Shellhaas RA. Phenobarbital and neonatal seizures affect cerebral oxygen metabolism: a near-infrared spectroscopy study. Pediatr Res. 2015; 78(1): 91–96.
14. Richardson DK, Corcoran JD, Escobar GJ, Lee SK. SNAP-II and SNAPPE-II: simplified newborn illness severity and mortality risk scores. J Pediatr. 2001; 138(1): 92–100.
15. Thompson CM, Puterman AS, Linley LL, et al. The value of a scoring system for hypoxic ischaemic encephalopathy in predicting neurodevelopmental outcome. Acta Paediatr. 1997; 86(7): 757–761.
16. Thorsen P, Jansen-van der Weide MC, Groenendaal F, et al. The Thompson encephalopathy score and short-term outcomes in asphyxiated newborns treated with therapeutic hypothermia. Pediatr Neurol. 2016; 60: 49–53. doi:10.1016/j.pediatrneurol.2016.03.014
17. Shankaran S, Laptook AR, Ehrenkranz RA, et al.; National Institute of Child Health and Human Development Neonatal Research Network. Whole-body hypothermia for neonates with hypoxic-ischemic encephalopathy. N Engl J Med. 2005; 353(15): 1574–1584.
18. Anders T, Emde R, Parmalee A. *A Manual of Standardized Terminology, Techniques and Criteria for the Scoring of States of Sleep and Wakefulness in Newborn Infants*. Los Angeles, CA: UCLA Brain Information Services; 1971.
19. Kirsch MR, Monahan K, Weng J, Redline S, Loparo KA. Entropy-based measures for quantifying sleep-stage transition dynamics: relationship to sleep fragmentation and daytime sleepiness. IEEE Trans Biomed Eng. 2012; 59(3): 787–796.
20. Walsh JL. A closed set of normal orthogonal functions. Am J Math. 1923; 45(1): 5–24.
21. Welch PD. The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. IEEE Trans Audio Electroacoust. 1967; AU-15: 70–73.
22. Johnson S, Moore T, Marlow N. Using the Bayley-III to assess neurodevelopmental delay: which cut-off should be used? Pediatr Res. 2014; 75(5): 670–674.
23. Palisano R, Rosenbaum P, Walter S, Russell D, Wood E, Galuppi B. Development and reliability of a system to classify gross motor function in children with cerebral palsy. Dev Med Child Neurol. 1997; 39(4): 214–223.
24. DuMouchel WH, O'Brien FL. *Integrating a Robust Option into a Multiple Regression Computing Environment*. Alexandria, VA: American Statistical Association; 1989.
25. Holland PW, Welsch RE. Robust regression using iteratively reweighted least-squares. Commun Stat Theory Methods. 1977; A6: 813–827.
26. Street JO, Carroll RJ, Ruppert D. A note on computing robust regression estimates via iteratively reweighted least squares. Am Stat. 1988; 42: 152–154.
27. Scher MS, Sun M, Steppe DA, Banks DL, Guthrie RD, Sclabassi RJ. Comparisons of EEG sleep state-specific spectral values between healthy full-term and preterm infants at comparable postconceptional ages. Sleep. 1994; 17(1): 47–51.
28. Freudigman KA, Thoman EB. Infant sleep during the first postnatal day: an opportunity for assessment of vulnerability. Pediatrics. 1993; 92(3): 373–379.
29. Shellhaas RA, Thelen BJ, Bapuraj JR, et al. Limited short-term prognostic utility of cerebral NIRS during neonatal therapeutic hypothermia. Neurology. 2013; 81(3): 249–255.

30. Peirano PD, Algarín CR. Sleep in brain development. Biol Res. 2007; 40(4): 471–478.
31. Tarullo AR, Balsam PD, Fifer WP. Sleep and Infant Learning. Infant Child Dev. 2011; 20(1): 35–46.
32. Scher MS, Steppe DA, Banks DL. Prediction of lower developmental performances of healthy neonates by neonatal EEG-sleep measures. Pediatr Neurol. 1996; 14(2): 137–144.
33. Scher MS, Loparo KA. Neonatal EEG/sleep state analyses: a complex phenotype of developmental neural plasticity. Dev Neurosci. 2009; 31(4): 259–275.
34. Scher MS, Johnson MW, Ludington SM, Loparo K. Physiologic brain dysmaturity in late preterm infants. Pediatr Res. 2011; 70(5): 524–528.
35. Beckwith L, Parmelee AH Jr. EEG patterns of preterm infants, home environment, and later IQ. Child Dev. 1986; 57(3): 777–789.
36. Burdjalov VF, Baumgart S, Spitzer AR. Cerebral function monitoring: a new scoring system for the evaluation of brain maturation in neonates. Pediatrics. 2003; 112(4): 855–861.
37. Hellström-Westas L, Rosen I, de Vries LS, Greisen G. Amplitude-integrated EEG classification and interpretation in preterm and term infants. Neoreviews. 2006; 7(2): e76–e87.
38. Csekő AJ, Bangó M, Lakatos P, Kárdási J, Pusztai L, Szabó M. Accuracy of amplitude-integrated electroencephalography in the prediction of neurodevelopmental outcome in asphyxiated infants receiving hypothermia treatment. Acta Paediatr. 2013; 102(7): 707–711.
39. Osredkar D, Toet MC, van Rooij LG, van Huffelen AC, Groenendaal F, de Vries LS. Sleep-wake cycling on amplitude-integrated electroencephalography in term newborns with hypoxic-ischemic encephalopathy. Pediatrics. 2005; 115(2): 327–332.
40. ter Horst HJ, van Olffen M, Remmelts HJ, de Vries H, Bos AF. The prognostic value of amplitude integrated EEG in neonatal sepsis and/or meningitis. Acta Paediatr. 2010; 99(2): 194–200.
41. Geva R, Yaron H, Kuint J. Neonatal sleep predicts attention orienting and distractibility. J Atten Disord. 2016; 20(2): 138–150.
42. Executive summary: Neonatal encephalopathy and neurologic outcome, second edition. Report of the American College of Obstetricians and Gynecologists' Task Force on Neonatal Encephalopathy. Obstet Gynecol. 2014; 123(4): 896–901.

ACKNOWLEDGMENTS

This research was supported by the National Institutes of Health (5K23HD068402) and the University of Michigan Ferrantino Investigator Award. The authors thank the research assistants, especially Diane White, RRT, CCRP, and Jamie Krinock, BS, the sleep technicians, especially Mark Kingen, RPSGT, and Laura Merley, RPSGT, and the Neonatal Intensive Care Unit follow-up clinic team, especially Ann Iatrow, for their important contributions to this project.

SUBMISSION & CORRESPONDENCE INFORMATION

Submitted for publication March, 2017
Submitted in final revised form July, 2017
Accepted for publication August, 2017
Address correspondence to: Renée Shellhaas, MD, MS, Pediatric Neurology, CS Mott Children's Hospital, room 12-733, 1540 E. Hospital Dr., Ann Arbor, MI 48109-4279, USA. Telephone: +734-936-4179; Fax: +734-763-7551; Email: shellhaa@med.umich.edu

This work was conducted at the C.S. Mott Children's Hospital, 1540 E. Hospital Dr., Ann Arbor, MI 48109, USA.

DISCLOSURE STATEMENT

RAS serves on the Board of the Child Neurology Society and on the Steering Committee of the Pediatric Epilepsy Research Consortium and receives royalties from UpToDate. RDC serves on the Board of the American Academy of Sleep Medicine and receives royalties from UpToDate. The following authors declare no conflicts of interest: Burns, Hassan, Carlson, and Barks.