ANNALS of Clinical and Translational Neurology
Open Access
AMERICAN NEUROLOGICAL ASSOCIATION

RESEARCH ARTICLE

Automated assessment of EEG background for neurodevelopmental prediction in neonatal encephalopathy

Micheline Lagacé[1,2], Saeed Montazeri[3], Daphne Kamino[4], Eva Mamak[5], Linh G. Ly[4,6], Cecil D. Hahn[2], Vann Chau[2], Sampsa Vanhatalo[3] & Emily W. Y. Tam[2,4]

[1]Faculty of Medicine, Clinician Investigator Program, University of British Columbia, Vancouver, British Columbia, Canada
[2]Division of Neurology, Department of Paediatrics, The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada
[3]BABA Center, Pediatric Research Center, Department of Clinical Neurophysiology, Children's Hospital, HUS imaging, HUS Diagnostic Center, University of Helsinki and Helsinki University Hospital, Helsinki, Finland
[4]Program in Neurosciences and Mental Health, SickKids Research Institute, Toronto, Ontario, Canada
[5]Department of Psychology, The Hospital for Sick Children, Toronto, Ontario, Canada
[6]Department of Paediatrics, The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada

Correspondence
Dr Emily W. Y. Tam, The Hospital for Sick Children, 555 University Ave, Toronto, ON M5G 1X8, tel: 416-813-6660, fax: 416-813-6709, E-mail: emily.tam@utoronto.ca

Received: 20 July 2024; Accepted: 6 October 2024

***Annals of Clinical and Translational Neurology* 2024; 11(12): 3267–3279**

doi: 10.1002/acn3.52233

Abstract

Objective: Assess the capacity of brain state of the newborn (BSN) to predict neurodevelopment outcomes in neonatal encephalopathy. **Methods**: Trends of BSN, a deep learning-based measure translating EEG background to a continuous trend, were studied from a three-channel montage long-term EEG monitoring from a prospective cohort of 92 infants with neonatal encephalopathy and neurodevelopmental outcomes assessed by Bayley Scales of Infant Development, 3rd edition (Bayley-III) at 18 months. Outcome prediction used categories "Severe impairment" (Bayley-III composite score ≤70 or death) or "Any impairment" (score ≤85 or death). **Results**: "Severe impairment" was predicted best for motor outcomes (24 h area under the curve (AUC) = 0.97), followed by cognitive (36 h AUC = 0.90), overall (24 h AUC = 0.84), and language (24 h AUC = 0.82). "Any impairment" was best predicted for motor outcomes (12 h AUC = 0.95), followed by cognitive (24 h AUC = 0.85), overall (12 h AUC = 0.75), and language (12 and 24 h AUC = 0.68). Optimal BSN cutoffs for outcome predictions evolved with the postnatal age. Low BSN scores reached a 100% positive prediction of poor outcomes at 24 h of age. **Interpretation**: BSN is an excellent predictor of adverse neurodevelopmental outcomes in survivors of neonatal encephalopathy after therapeutic hypothermia, even at 24 h of life. The trend provides a fully automated, objective, quantified, and reliable interpretation of EEG background. The high temporal resolution supports continuous bedside brain assessment and early prognostication during the initial dynamic recovery phase.

Introduction

Neonatal encephalopathy (NE) is a significant concern in the neonatal period[1] with hypoxic–ischemic encephalopathy (HIE) being the principal etiology.[1] Very early prognostication of neurodevelopmental outcomes is essential for guiding optimal neurocritical care in neonatal intensive care units. It is typically done by incorporating information from multiple modalities.[2–5]

Continuous brain monitoring with scalp-recorded electroencephalography (EEG) during the first days of life is crucial for neuroprognostication, which is best done by assessing the recovery of normal spontaneous cortical activity (also known as "EEG background (activity)"), including re-appearance of sleep–wake cycling.[2,6–8] It is well established that a persisting severely abnormal EEG background predicts adverse neurodevelopmental outcomes, while early recovery of EEG background and emergence of sleep–wake cycles support favorable outcomes.[2,6–8] Unfortunately, reliable bedside assessments of these key EEG characteristics present a challenge.[9] A time-compressed trend called amplitude-integrated EEG

(aEEG) is used to facilitate bedside review.[7,10] However, aEEG interpretation is still subjective and requires substantial expertise.[2,6,9] Thus, there remains a need for objective, quantified and preferably automated bedside assessment of neonatal EEG and aEEG monitoring.

A fully automated trend measure, brain state of the newborn (BSN), was recently developed for a quantified and objective assessment of EEG background through brain recovery in NE during the first days of life.[11] It is an open-access, validated, and reliable tool that could facilitate and harmonize neonatal EEG monitoring for clinical and research purposes.[11] BSN can be produced with a single bipolar channel, but including additional channels was shown to provide more stable BSN trends, which can be displayed on a bedside monitor with adjustable time resolutions (minutes to hours).[11] The prior BSN development study reported early diverging BSN trends in infants with severely abnormal outcomes.[11] Using clinical outcome categories at 4 years of age (favorable, cerebral palsy, cerebral palsy with epilepsy, and death), BSN showed accurate outcome predictions during early hours after birth asphyxia.[12] BSN also distinguished between severity of encephalopathy and predicted abnormal neurodevelopmental outcomes at 2 years.[13] It is not well established, however, whether BSN could provide very early prediction of more detailed neurodevelopmental outcomes, such as differential predictions of the neurodevelopmental domains (motor, cognitive and language) that are commonly assessed in neuropsychological batteries.

We aimed to assess how the early BSN trends can predict domain-specific neurodevelopmental outcomes in NE in a larger prospectively cohort. BSN trends were computed offline from EEG data of an observational prospective cohort of infants with NE, and they were used to predict the domain-specific neurodevelopmental outcomes at 18 months. We hypothesized that low or slow-to-improve BSN trends in NE undergoing therapeutic hypothermia would predict adverse neurodevelopmental outcomes, while high or fast-to-recover BSN trends would predict favorable neurodevelopmental outcomes.

Methods

Study design and participants

A secondary analysis of a cohort of infants with NE used to assess how BSN can predict domain-specific neurodevelopmental outcomes. The infants were recruited at a single outborn tertiary center, the Hospital for Sick Children in Toronto, between 2014 and 2019. This observational cohort was created to assess the role of dysglycemia during NE.[14] The dysglycemia information was not used for this specific project. Infants were eligible if they presented with NE, defined by abnormal consciousness with either neonatal seizures or abnormalities in tone or reflexes.[15] Exclusion criteria included PMA less than 36 weeks, birth weight below 1500 g, congenital malformations, inborn errors of metabolism, or congenital infections. Infants who could not be recruited by 6 h of life were excluded.

The infants received standard of care for NE, including therapeutic hypothermia. Therapeutic hypothermia status, such as rewarmed early, was not an exclusion criterion to improve generalizability. Demographic and clinical data were collected. Informed consent was obtained from parents or legal guardians following a protocol approved by the Hospital for Sick Children's Research Ethics Board. Apart from EEG signals, study data were managed using research electronic data capture (REDCap, Vanderbilt University, Tennessee)[16] hosted at the Hospital for Sick Children. STROBE checklist for observational cohort studies was respected for this project.

EEG monitoring

Continuous EEG monitoring was started as soon as possible and continued for at least 48 h. We recorded the EEG signals at 200, 250, or 256 Hz[17,18] using either Stellate Harmonie or Xltek Brain Monitor ICU video-EEG systems (Natus Neurology, Oakville, Ontario, Canada). Using Elefix electrode paste, 11 or 20 grass gold cup surface electrodes were attached individually following the International 10–20 system using Pz as reference, evolving from the 10–20 system modified for neonates to the full 10–20 system before the end of the study period. The EEG signals were saved in the generic European Data Format (.edf) using a three-channel montage that was both available in the whole cohort and that closely corresponded to the limited-channel aEEG monitoring[9]: two hemispheric signals (Fp1-C3 and Fp2-C4) and one cross-cerebral signal (C3-C4).

Generating the BSN trends

After re-montaging and EDF conversion, the EEG signals were submitted to a computational server (www.babacloud.fi) for a fully automated analysis pipeline,[11] which outputs channel-wise BSN values and an automated artifact classification[19] for each second of the EEG file. The Babacloud server operates fully automatically, and it only requires the user to upload the EEG file using a simple web interface, followed by downloading the analysis results for each EEG file.[20] As an alternative to the web browser, the analysis pipeline can be

also used directly via Python-based API script. The training and external validation of the analysis algorithm was previously explained.[11] The BSN algorithm converts the raw EEG signals into a continuous score from 0 to 100, where 0 corresponds to a fully inactive EEG, and 100 represents continuous EEG with sleep–wake cycles.[11] Sleep–wake cycling are presented in BSN by high score with relatively low amplitude rhythmic fluctuations going downward for quiet sleep and up toward 100 for active sleep and wake state.[11] The BSN outputs reported in this work present average BSN values from the accepted bipolar derivations (see Appendix p. 3) and their confidence intervals, jointly visualized as *BSN trends* (Fig. 1). During this post-processing, artifacts and seizures were automatically removed (see Appendix p. 3). We visually inspected the individual BSN trends with their corresponding aEEG trends for an *ad hoc* quality assurance and to ensure sufficient clinical explainability[21] of the BSN findings.

Neurodevelopmental assessment

The survivors of NE had a neurodevelopmental assessment at 18 months including Bayley Scales of Infant Development, 3rd edition (Bayley-III), administered by a trained assessor supervised by a registered psychologist. Reported composite scores included motor, cognitive, and language. Participants who were unable to engage in testing were assigned specific scores[14] following literature.[22] A diagnosis of cerebral palsy (CP) was assigned a composite score for motor, cognitive, and language of 70 (−2 standard deviation (SD) on Bayley-III) if their CP severity was Gross Motor Function Classification System (GMFCS) III ($n = 1$) and 45 (−3SD) if their severity was GMFCS IV or V ($n = 1$). Participants unable to engage in testing received a motor composite score of 100 if they walked independently by 14 months ($n = 3$) and a language composite score of 70 if they were non-verbal ($n = 2$).

Outcome grouping

We studied two alternatives to define adverse neurodevelopmental outcomes. The more conservative approach ("severe impairment") used death or a Bayley-III composite score ≤70 corresponding to −2 standard deviation (SD) and moderate–severe impairment.[23] The more inclusive approach ("any impairment") used death or a Bayley-III composite score ≤85 corresponding to −1SD and mild or at risk of impairments.[23] Overall adverse neurodevelopmental outcomes were defined as a below-threshold score in at least one of the composite scores or death prior to follow-up.

Statistical analysis

Descriptive statistics were performed on the demographics and clinical information. Receiver operating characteristic (ROC) curves were calculated for BSN values at 12, 24, 36, and 48 h of life. The BSN-based prediction in the ROC curves was considered poor (area under the curve (AUC) = 0.50–0.69), acceptable (AUC = 0.70–0.79), excellent (AUC = 0.80–0.89), or outstanding (AUC >0.90).[24] The evolution of the optimal BSN cutoff levels and their corresponding positive predictive value (PPV) and negative predictive values (NPV) were analyzed as a function of postnatal age and presented in 3D plots. To produce the ROC curves and the 3D plots with PPVs and NPVs, the median BSN value over the specific time, preceding, and following hour were used (e.g., BSN at 12 h = median [BSN at 11 h, BSN at 12 h, and BSN at 13 h]). Statistical analysis was done using Matlab version R2022b and STATA 18.0 software (Stata Corp., College Station, TX). The BSN outputs and figures were visualized with Matlab version R2022b.

Role of funding source

None of the funders had a role in the design and conduct of the study, collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

Results

Of the 100 infants from the cohort involved in this project, six participants were removed from the analysis due to absence of Bayley-III evaluation (withdrew ($n = 1$), lost to follow-up ($n = 4$), and missed follow-up ($n = 1$)), one because a suspected genetic condition, and another due to technically inadequate EEG. The final analyses included 92 infants, 61% males, who underwent neonatal continuous EEG and neurodevelopmental evaluation at 18 months.

The demographic and clinical characteristics are presented in Table 1. We considered the cohort to represent neonates receiving therapeutic hypothermia in NE since only one infant received no therapeutic hypothermia. Continuous EEG was connected on average by 17.62 ± 8.35 h. Only 26 infants were connected before 12 h of life. Neonatal seizures were reported in 49 infants (53%), and 50 infants (54%) received at least one anti-seizure medication (ASM) before or during EEG recording. At least one analgesic or sedative agent was administered to 97% of infants by 24 h of life, principally morphine infusion (84%), but bolus of morphine,

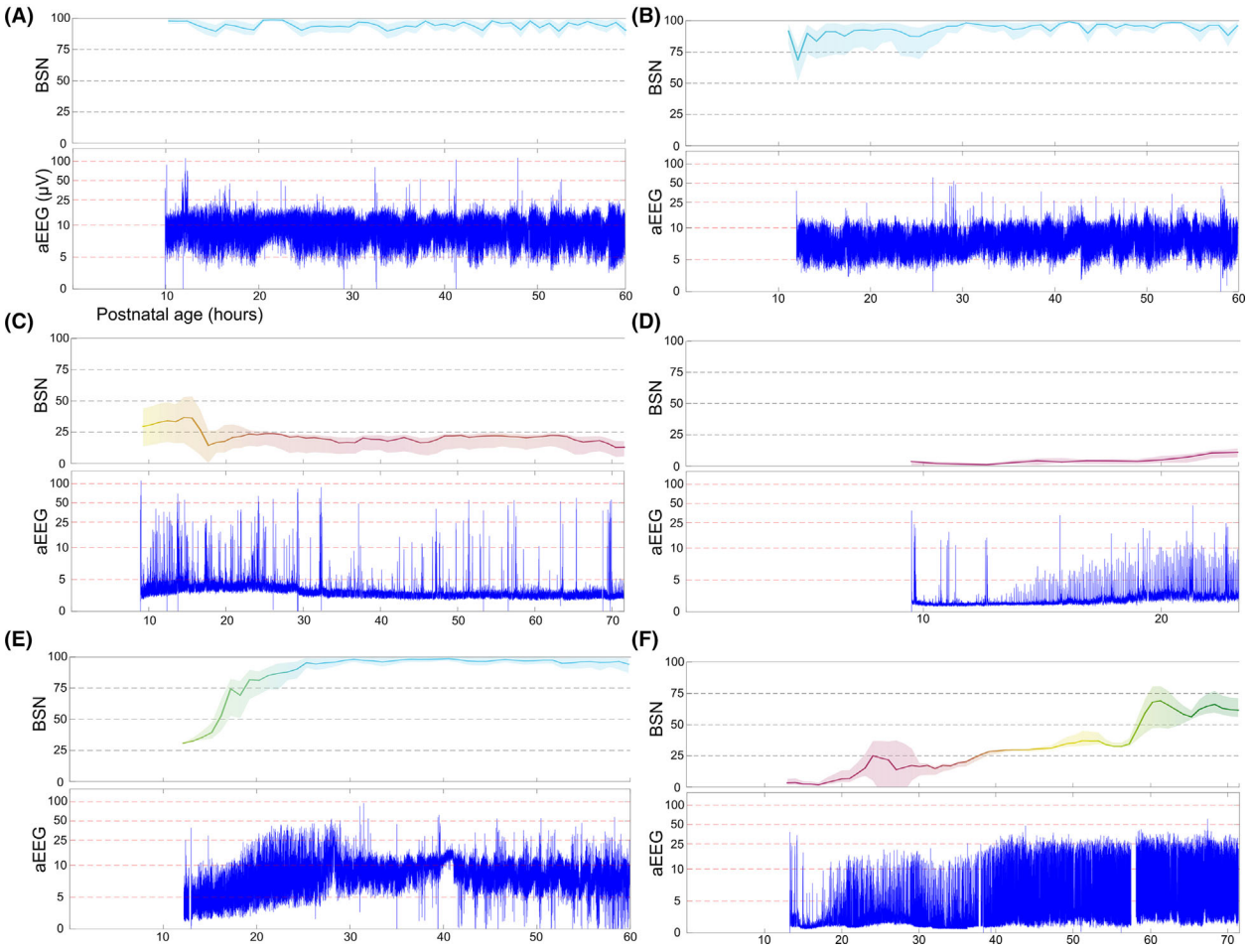

Figure 1. Illustrative examples of brain state of the newborn (BSN) trends with their corresponding amplitude-integrated EEG (aEEG) trends. The color in BSN trends is coding for the level (0–100), while the shade around the BSN trends shows confidence interval. Observational descriptive terms suggest 24–36 h to part early or quick versus late or slow, while >80–90s is offered for high BSN score and <30 for low. aEEGs shown correspond to the signal from channels C3–C4. (A) High BSN scores reflect a good background and normal neurodevelopmental outcomes. The relatively low amplitude rhythmic fluctuations of the high score BSN trends represent the sleep–wake cycling, with the BSN trends going downward for quiet sleep and going up towward 100 for active sleep or wake state. aEEG with continuous normal voltage and sleep–wake cycling. (B) BSN scores with some initial non-rhythmic variability but remaining high, reflecting a good background and normal neurodevelopmental outcomes. The low amplitude rhythmic fluctuation of sleep–wake cycling becomes more evident in the second half of the recording. aEEG with continuous normal voltage and sleep–wake cycling. (C) Low BSN scores reflect poor background and severe neurodevelopmental impairments. aEEG with burst suppression. (D) Low BSN scores reflecting poor background in a participant with neonatal death. aEEG with burst suppression. (E) BSN scores with early improvement and normal neurodevelopmental outcomes. aEEG with initially low voltage with evolution to discontinuous normal voltage and finally continuous normal voltage with sleep–wake cycle. (F) BSN scores with slower improvement and adverse neurodevelopmental outcomes. aEEG with burst suppression initially, which evolved to discontinuous normal voltage.

fentanyl, or dexmedetomidine were also used. No infant remained naive to neuroactive medication, including ASMs, sedation, and analgesia.

All the participants had a neurodevelopmental assessment between 17 and 25 months (mean: 18.95 ± 1.69). Neurodevelopmental outcomes are presented in Table 2. All participants had composite scores completed for the three domains of the Bayley-III, except for one infant who could not be tested for language due to a language barrier. There were nine neonatal deaths and one post-neonatal death.

The recoveries of the individual BSN trends were visually compared to neurodevelopmental outcomes. Visual analysis of BSN trends was essential to understand patterns and identify outliers. However, the time-linked ROC analyses, PPVs, and NPVs offer the most objective analysis of BSN prediction capacity. Hence, observational descriptive terms suggest 24–36 h to part early or quick

Table 1. Subject demographics with clinical and therapeutic characteristics.

Variables	Entire cohort (*N* = 92)	Overall adverse neurodevelopmental outcomes	
		Severe impairment (*N* = 21)	Any impairment (*N* = 40)
Sex	Male 56/92 (61%)	Male 13/21 (62%)	Male 25/40 (63%)
Post-menstrual age, weeks	39.6 ± 1.3	39.4 ± 1.3	39.6 ± 1.2
Birth weight, g	3411 ± 510	3343 ± 436	3442 ± 523
Umbilical cord gas			
pH	7.00 ± 0.17	6.97 ± 0.22	7.01 ± 0.20
Base excess	−15.21 ± 6.95)	−17.14 ± 7.34	−15.63 ± 6.78
APGAR at 5 min	4.0 (3.0–6.0)	4.0 (1.0–6.0)	3.5 (1.5–6.0)
Sarnat score			
Mild	2/92 (2%)	1/21 (5%)	1/40 (3%)
Moderate	77/92 (84%)	7/21 (33%)	26/40 (65%)
Severe	13/92 (14%)	12/21 (62%)	13/40 (33%)
Therapeutic hypothermia			
Complete (72 h)	86/92 (93%)	19/21 (90%)	37/40 (92.5%)
Partial (<72 h)	5/92 (5%)	2/21 (10%)	3/40 (7.5%)
None	1/92 (1%)	0/21 (0%)	0/40 (0%)
Age at the start of EEG, h	17.6 ± 8.4 (range: 6.1–44.8)	14.5 ± 7.0 (range: 6.1–31.0)	15.1 ± 7.0 (range: 6.1–35.6)
Infants with neonatal seizure	49/92 (53%)	17/21 (81%)	28/40 (70%)
Antiseizure medication,[a] number of agents	1.0 ± 1.2	2.0 ± 1.4	1.5 ± 1.5
Analgesic and sedative medication,[a] number of agents	1.3 ± 0.5	1.4 ± 0.7	1.3 ± 0.6
Neuroactive medication,[a] number of agents	2.4 ± 1.4	3.4 ± 1.7	2.8 ± 1.6
Deceased prior to 18 months follow-up	10/92 (11%)	10/21 (48%)	10/40 (25)
Diagnosis cerebral palsy at 18 months follow-up	6/92 (7%)	6/21 (55%)	6/40 (20%)
Maternal education			
Less than high school	2/92 (2%)	0/21 (0%)	1/40 (3%)
High school diploma	4/92 (4%)	0/21 (0%)	0/40 (0%)
College or specialized training	21/92 (23%)	2/21 (10%)	8/40 (20%)
Undergraduate university degree	8/92 (9%)	1/21 (5%)	3/40 (8%)
Graduate degree	5/92 (5%)	1/21 (5%)	1/40 (3%)
Not available	52/92 (57%)	17/21 (81%)	27/40 (68%)
Primary language spoken at home			
English	36/92 (40%)	3/21 (14%)	12/40 (30%)
Other than English	5/92 (5%)	1/21 (5%)	2/40 (5%)
Not available	51/92 (55%)	17/21 (81%)	26/40 (65%)

Overall adverse neurodevelopmental outcomes of Bayley-III when "severe impairment" corresponds to at least one Bayley-III composite score ≤70 or death while "any impairment" corresponds to at least one composite score ≤85 or death. Neuroactive medication includes anti-seizure medication, sedation, or analgesia. Data are *n*/*N* (%), mean ± SD, or median (IQR).

[a]Prior to or during EEG recording.

versus late or slow, while >80–90s is offered for high BSN score and <30 for low. The most frequent observation was that early and sustained high BSN levels with sleep–wake cycling were observed in neonates with favorable outcomes (Fig. 1A,B). Low BSN levels were observed in neonates with adverse outcomes (Fig. 1C,D), including moderate–severe neurodevelopmental impairments or neonatal death. Early improvement in BSN was observed in neonates with favorable outcomes (Fig. 1E), while slow and late BSN improvement was observed with adverse outcomes (Fig. 1F).

Ten infants were identified with higher BSN levels that abruptly dropped to lower BSN levels, showing a subsequent BSN recovery and reassuring outcomes (six normal and four mild impairment) (see Appendix p. 4, Fig. S1A–C for illustrative examples). One participant with "severe impairment" had a BSN drop with a plateau before slow recovery (Fig. S1D). Hence, the later improved BSN trend, with their preceding BSN if the decrease was at an older age, appeared to predict respective outcomes as normal or mild impairments, while a drop with a persisting lower or slow-to-recover BSN score was seen with adverse

Table 2. Domain-specific neurodevelopmental outcomes at 18 months.

Bayley-III (composite scores)	Adverse neurodevelopmental outcomes (N = 92)	
	Severe impairment	Any impairment
Overall	21 (23%)	40 (43%)
Motor	17 (18%)	23 (25%)
Cognitive	18 (20%)	25 (27%)
Language	20 (21%)	37 (40%)

Adverse neurodevelopmental outcomes for each composite score of Bayley-III when "severe impairment" corresponds to a Bayley-III composite score ≤70 or death while "any impairment" corresponds to a composite score ≤85 or death.

outcomes. The abrupt drops in BSN were at times related to a clinical change, such as seizures and ASM; however, some BSN drops were not associated with an identifiable clinical event. The same proportion of participants with and without a clinical correlation for the BSN drop and with either normal or mild developmental concerns were detected precluding the identification of a pattern.

We also identified infants with an apparent mismatch between high initial BSN levels (BSN ~90 before 24 h) and adverse neurodevelopmental outcomes (see Appendix p. 5, Fig. S2A–F). One such infant died of sudden infant death syndrome (SIDS) after the neonatal period (Fig. S2A). The infant was kept in the analyses because it was impossible to exclude the role of HIE in later SIDS.[25,26] Four other outliers showed high early BSN levels and moderate-to-severe impairment in at least one development domain, but all included language (Fig. S2C–F). An infant had high BSN levels with moderate-to-severe impairment in all domains, but the EEG was started after 30 h (Fig. S2B). This disproportional involvement of language (14/15 participants) was more marked when including Bayley-III composite score 71–85 with isolated language impairment in 9/15 participants. Only one participant was confirmed to have a primary language other than English, but primary language was unavailable for 10 out of 19 infants with language impairment and high BSN trend. Maternal education, a marker for socioeconomic status, was unavailable for most participants (52/92).

Further analysis with ROC curves showed how prediction by BSN levels changes over time and between outcome groups (Fig. 2). BSN was excellent to outstanding at predicting "severe impairment" for the four postnatal ages assessed. Time-linked ROC analyses showed AUCs of 0.80 to 0.97 for severe overall impairment. BSN best-predicted motor outcomes at 24 h of life (AUC = 0.97), followed by overall outcomes (AUC = 0.84) and language (AUC = 0.82) outcomes, while cognitive outcomes were maximally predicted at 36 h (AUC = 0.90). The 3D graphs of PPV and NPV of BSN score per postnatal age (Fig. 3) showed that the "severe impairment" group required a lower BSN score to reach the same PPV than the "any impairment" group; at 24 h overall outcome had a PPV of 100% for both groups, but the required BSN score was 15 for "severe impairment" and 40 for "any impairment" (see Appendix p. 6, Tables S1 and S2). For the "any impairment" group, the prediction capability of BSN was more variable, with AUCs between outstanding and poor (Fig. 2). The prediction was the strongest for motor outcome at 12 h of life (AUC = 0.95), followed by overall outcomes (AUC = 0.75), and language (AUC = 0.68), while cognitive outcomes were maximally predicted at 24 h (AUC = 0.85). ROC curves of adverse neurodevelopmental outcomes in function of BSN score for our two outcome groups showed that adverse outcomes were better predicted using ≤70 as the composite score limit at the four postnatal ages assessed.

The three-dimensional (3D) relationships (BSN-level × time × PPV/NPV) were assessed using time-varying analysis of PPVs and NPVs through all possible BSN cut-off levels using our two neurodevelopmental outcome groups (Fig. 3). This approach depicted the "optimal BSN" cutoff levels with the highest possible PPV/NPV at each point in time. The maximal PPV and NPV at four postnatal ages with their corresponding BSN scores are reported for the two outcome groups in Tables S1 and S2. As example, using "severe impairment," the maximal PPV at 24 h was 100% for all domains for BSN scores of 15, while maximal NPV was 93% for overall, motor 100%, cognitive 96%, and language 93% for BSN scores of 60. For all developmental spheres in the two groups, the maximal PPVs declined in the first 24 h of life before showing sustained high PPV. After 24 h of life, there was a general trend where the best prediction was obtained with higher BSN cut-off levels as the infants grew older. Thus, with "severe impairment," motor maximal PPV was 86% at 12 h but 100% at 24, 36, and 48 h, but the respective BSN scores were 45, 15, 60, and 80. The increasing optimal BSN cut-off was most apparent for the overall (PPV) and motor (PPV and NPV) outcomes. Comparison between outcome domains shows that the best PPV of cognitive or language outcomes comes from much lower BSN levels than overall and motor outcomes; however, NPV of the cognitive and language outcomes is optimal at high BSN levels. Using "any impairment" at 24 h, maximal motor PPV was 100% for BSN score of 40, while cognitive maximal PPV was 100% for BSN score of 15 and maximal NPV was 94% for a BSN score 60. The relatively wide difference in BSN levels between PPV and NPV for cognitive and language outcomes is also reflected

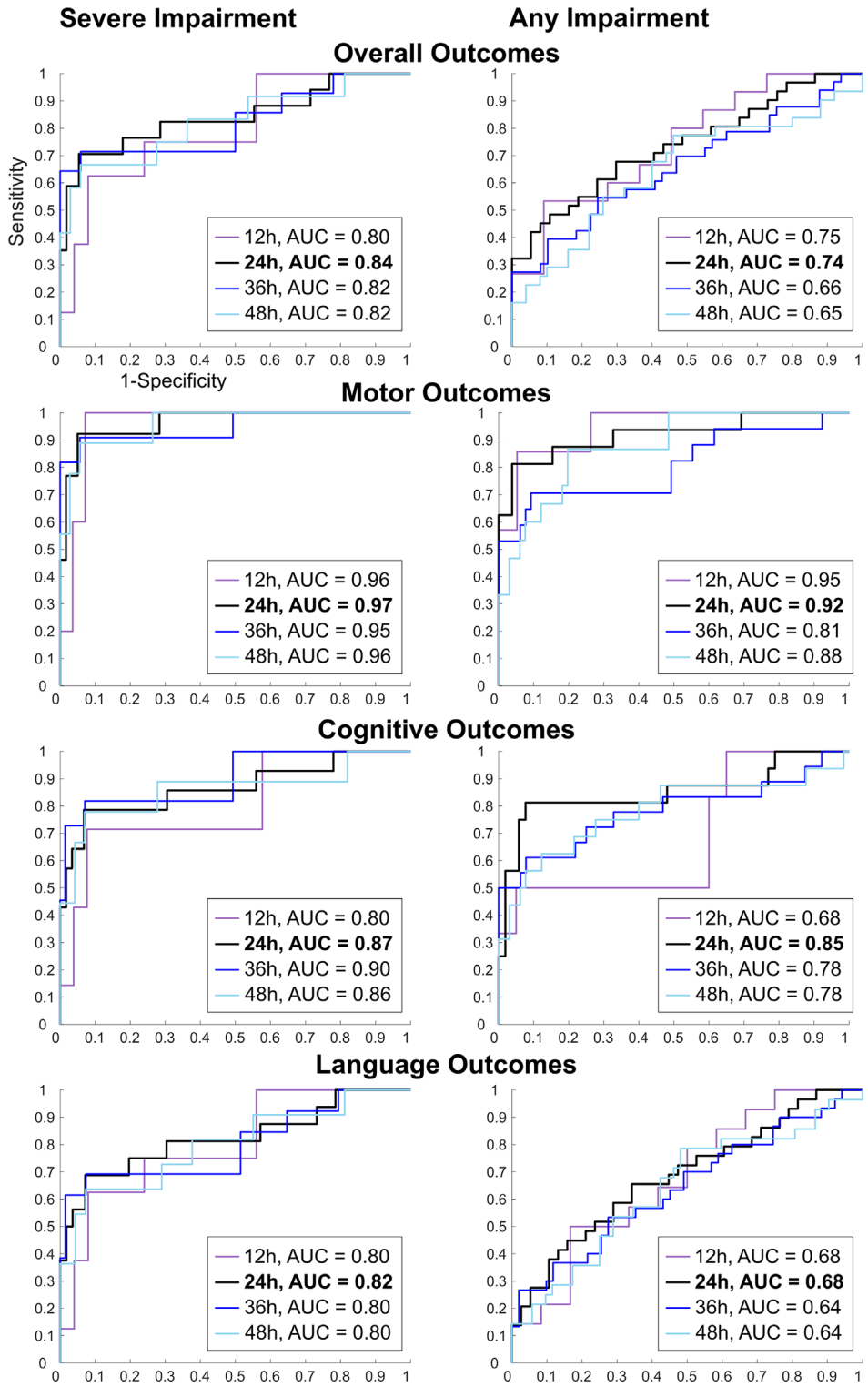

Figure 2. Receiver operating characteristic (ROC) curves for neurodevelopmental outcome predictions. Each ROC curve depicts adverse neurodevelopmental outcomes in function of brain state of the newborn (BSN) at four different postnatal time points (12, 24, 36, and 48 h). The "severe impairment" group corresponds to a Bayley-III composite score ≤70 or death, while the "any impairment" group corresponds to a composite score ≤85 or death. AUC, area under the curve.

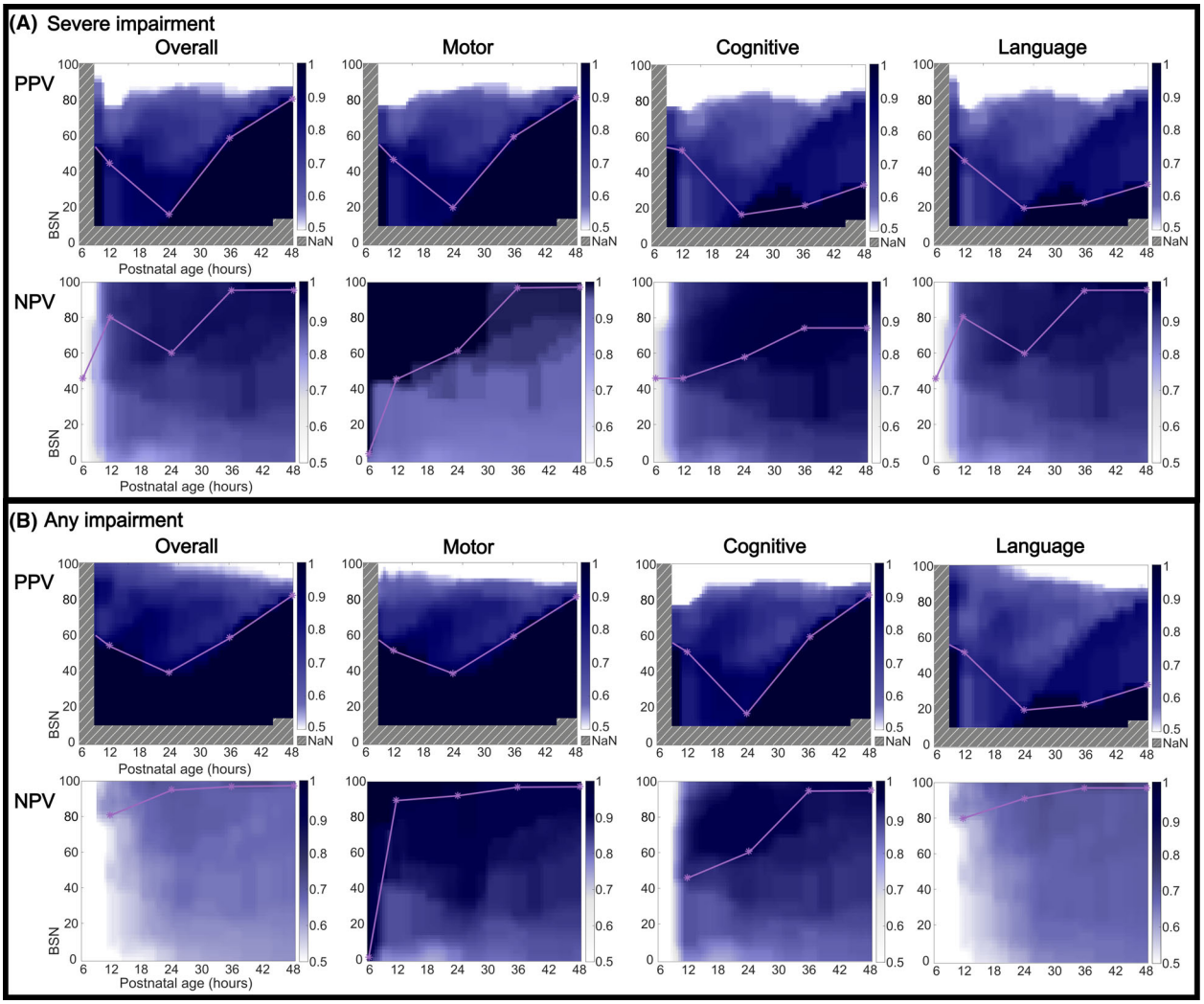

Figure 3. Change in outcome predictions by different brain state of the newborn (BSN) levels during the first 2 days of life. These three-dimensional graphs show how choosing different BSN cutoff values would affect positive predictive value (PPV) and negative predictive value (NPV). The light purple line follows the optimal BSN level at each hour, showing what BSN level would have maximal PPV or NPV, respectively. The color scales depict PPV and NPV levels scaled to the same color schemes across the whole figure (see color bar on the right side); darker colors indicate higher predictive value. The "severe impairment" group corresponds to a Bayley-III composite score ≤70 or death, while the "Any impairment" group corresponds to a composite score ≤85 or death. NaN, data not available for calculation.

in the lower corresponding AUC values. With any language impairment, the PPV of 100% at 24 h corresponded to BSN 20, but the NPV of 70% had a BSN score of 90 while the AUC was 0.68. PPV of cognitive outcomes was influenced by the Bayley-III cutoff used; PPV at 36 h was 100% for both cut-offs, but BSN score was 20 for "severe impairment" and 60 for "any impairment."

A diagnosis of CP was given to six participants, all with adverse motor outcomes regardless of the cutoff used. Using PPVs, adverse motor outcomes could be predicted at 100% at 24 h in two participants (BSN <15) and two others at 36 h (BSN <60). A participant with CP was an outlier (Fig. S2B) and the EEG of the sixth participant was started too late (>48 h) to assess BSN prediction.

Our findings suggest favorable neurodevelopment when BSN levels are persistently over 60 by 36–48 h of life; however, BSN may offer strong prediction already by 24 h. Tables S3 and S4 (see Appendix p. 7) present PPVs and NPVs for specific BSN scores and postnatal ages for the two outcome groups. Hence, using "severe impairment," a BSN score of 60 at 36 h gives NPVs above 90% for all domains with a PPV of 100% for overall, 83% for motor, and 73% for cognitive and language. However, a

BSN of 60 at 24 h also gives NPVs above 90% for all domains, but with a PPV of 73% for overall and motor, and 67% for cognitive and language. The neurodevelopmental prediction by BSN at 24 h was outstanding for motor and excellent for cognitive regardless of the Bayley-III cutoff used (Fig. 2). The prediction of overall outcomes went from excellent to acceptable with Bayley-III ≤70 versus ≤85. Language prediction was more affected by the Bayley-III score used, going from excellent with ≤70 to poor with ≤85. BSN offered high PPV and NPV at 24 h (Fig. 3). The maximal PPV at 24 h used lower BSN (BSN ≤15) for overall and motor "severe impairment" than for "any impairment" (BSN ≤40). At 24 h, maximal NPV was around BSN 60 for "severe impairment" but 90 for "any impairment" for overall, motor, and language outcomes.

Discussion

Our results show that BSN trends at 24 h can predict later domain-specific neurodevelopmental outcomes in NE receiving therapeutic hypothermia. BSN trends offer effective, explainable, and intuitive displays of the recovery of cortical activity during the first days of life in NE. These findings concord with the extensive literature where visually analyzed aEEG[27] and EEG background shows predictions of later outcomes.[6,7,10] We extend those findings by demonstrating that the analysis of EEG and aEEG signals can be fully automated to provide an objective, continuous measure of the EEG background activity. Moreover, our results show that such continuous trend measures can provide accurate outcome predictions as early as the first day of life in hypothermia-treated infants, which is a significant improvement to the visual EEG analyses that become predictive by 36 h or more.[7,10,27–29]

The changes in specificity and sensitivity of BSN-based predictions are compatible with the previous literature using visual observations of EEG and aEEG trends[7,10,12,27]; sensitivity decreases and specificity increases over the early neonatal period. Compared to the visual EEG and aEEG assessments, BSN shows an earlier prediction, with peak AUCs between 12 and 36 h for each developmental domain. This is a clinically significant improvement provided by BSN, as therapeutic hypothermia is commonly known to delay prognostication by conventional EEG or aEEG from 24–36 h to 48–72 h.[7,27–29] A more recent meta-analysis, only using studies with follow-ups of 18 months or more, identified the highest diagnosis odd ratio of aEEG at 36 h.[10]

Our analyses demonstrate robust dynamics in the BSN-based prediction, supporting ways to customize and optimize outcome predictions for individual cases or trial designs. For instance, BSN can have a PPV of 100% by 24 h of life when using a very low BSN score (BSN ≤15). This degree of precision with BSN is due to its continuous value instead of the conventionally used discrete classes of background, which reported PPVs of 66% at 24 h, 85% at 48 h, and 89% at 72 h for persistently abnormal aEEG and adverse neurodevelopment outcomes.[27] Without age division, a meta-analysis reported an AUC of 0.78 for unfavorable neurological outcomes using aEEG background patterns, while EEG was 0.88.[30] In a meta-analysis of EEG in therapeutic hypothermia, the sensitivity and specificity of burst suppression, low voltage, and flat trace pattern, respectively varied between 0.84 to 0.87 and 0.60 to 0.94.[7] A detailed comparison of the prior literature[7,10,27,30] to the present study is challenged by the differences in the EEG measures used in each study. However, a combination of discrete background patterns and sleep–wake cycle in the aEEG trend was found to predict adverse 2-year neurodevelopmental outcomes with Bayley <70 as the threshold, with AUCs ranging from 0.89–0.91 for 12–24 h to 0.90–0.92 for 24–36 h.[3]

Prognostication of neurodevelopmental outcomes following therapeutic hypothermia in NE is multimodal using tools outside EEG. Moderate-to-severe modified Sarnat score before 6 h of life showed an AUC of 0.72 for adverse neurodevelopment outcomes at 2 years.[31] Their outcome definition was closer but more inclusive than our overall "any impairment" with an AUC of 0.74 at 24 h.[31] Using 2 years outcomes, visual evoked potential showed a PPV of 91% for adverse outcomes.[32] However, NPV was only 58%.[32] A meta-analysis reported an AUC of 0.88 for MRI within 2 weeks of age to predict unfavorable neurological outcomes.[30] Since, it was brought forward that MRI predictive capacity is reduced in the absence of severe brain injury and cannot accurately discriminate the degree of neurodevelopmental impairments.[33]

Adverse neurodevelopmental outcomes have diverse definitions, including variability in standardized tests and thresholds used, as well as the practical implications of neurodevelopmental compromise in different walks of life.[7,23,34] Our work used both common cutoffs for adverse neurodevelopmental outcomes to facilitate comparison with literature. The findings show the expected trade-offs: The BSN-based prediction was excellent to outstanding in predicting adverse neurodevelopmental outcomes when using death or Bayley-III composite score ≤70, but it was outstanding to poor when using ≤85. Higher specificity with less false positives may be clinically useful, if BSN is used for supporting clinical management of NE, including discussions around goals of care and redirection.

Unfortunately, most studies do not assess the predictive capacity by developmental domains but use composite

adverse outcomes. Regarding domain-specific prediction, our ROC analyses presented that BSN trends show best prediction of motor outcomes and lowest prediction of language outcomes. However, our ad hoc observation of reassuring BSN trends in infants with poor language outcomes was compatible with the idea that language might be substantially modified by later environmental factors beyond the effects of early brain injury, obviously not predictable in the newborn EEG data.[35,36] The literature increasingly reports the role of socioeconomic status and environment on outcomes after neonatal brain injury,[35,37] with language and cognitive being the domains the most influenced by these modifying factors.[35,36] Also, with multiculturalism, bilingualism and English as a secondary language need to be considered during testing and analysis.[38,39] Furthermore, the reported overall prevalence of neurodevelopmental impairment in the general pediatric population, 7.6% to 17.8%,[40,41] coincides closely with the 7% proportion of apparent prediction failures after therapeutic hypothermia in NE.[27]

Despite measuring the same phenomenon, EEG background, the advantage of BSN[12] to aEEG[27] or EEG.[6,7,10] is likely attributed to the key differences between visual analysis and the BSN construct.[11] Visual EEG and aEEG analyses are based on identifying discrete categories (e.i. inactive, burst-suppression, and continuous), forming a continuum from worst (inactive) to best (fully active). The discrete nature of EEG background categories overcomes constraints in human visual perception, which contradicts extensive empirical data and clinical experience on a confluence between neonatal brain states. With BSN, such artificial constraints are removed by replacing the categories with a continuous value, increasing the "depth resolution" of EEG background activity. In practice, this becomes important for outcome predictions because infants with NE may show a prolonged recovery with an almost normalized EEG activity. Such EEG activity may look partly discontinuous and abnormal to varying degrees. However, its visual classification is often confusing[5–7,27] falling between different categories. In the aEEG nomenclature, it can be called discontinuous normal voltage[2,5,6] but the classification is ambiguous and subject to confounders. BSN also allows the inclusion of a richer repertoire of EEG and aEEG information content that can be visually perceived. Thus, we propose that the good performance of BSN in early prediction comes from its continuous measure that spans through the gray zones in visual classifications (BSN levels between 50 and 80).

Limitations

There are limitations in the present study. First, we were unable to assess infants before 12 h of life because monitoring was delayed from being an outborn center. Second, minutes with seizures were removed, respecting the design of BSN focusing on background analysis.[11] Visual or automated seizure detection is needed to supplement BSN in potential clinical use.[8,9] Third, any neuroactive agents, given to all infants in our cohort, are known to affect EEG and aEEG background,[29] and hence BSN levels; these effects will require further detailed, time-linked analyses, due to possible confounding effects. However, BSN still demonstrated excellent predictive capacity despite using neuroactive agents supporting its generalizability to its relevant population; infants with encephalopathy receiving various neuroactive agents. Fourth, longer follow-up periods could allow more complex neurodevelopmental evaluations. However, such work would come at the cost of confounding the neonatal effects with other early life effects that the EEG or aEEG monitoring cannot explain during the first days of life. Fifth, visual inspection of the BSN trends was not blinded to outcomes because understanding the interaction between BSN trends and neurodevelopment was intrinsic to this project; knowledge of outcomes was necessary to identify if a BSN trend was concordant or not with expectations but may have introduced some bias. Finally, diagnosis of autism spectrum disorder was not collected and only limited information about primary language spoken at home and maternal education were available precluding their assessment for possible confounding impacts on neurodevelopmental outcomes.

Conclusion

BSN may provide excellent prediction of domain specific adverse neurodevelopmental outcomes in survivors of NE after therapeutic hypothermia. Low or slowly improving BSN trend links to adverse neurodevelopmental outcomes, while high or fast recovering BSN trend links to favorable neurodevelopmental outcomes. Prediction of more severe neurodevelopmental outcomes is more specific, providing critical information for management. The principal advantage of BSN over aEEG is its earlier predictive capability, which is maximal already at 24–36 h of life. Other benefits of BSN against aEEG or EEG include its objectivity, quantifiability, and reliability from being an open-access automatic algorithm, which also alleviates the need for expert interpretation. In summary, BSN is a promising open-access algorithm that interprets EEG background in NE and offers bedside neuroprognostication.

Acknowledgements

The authors thank the families and children who have participated in this study. The authors also thank the

NOGIN research team, including the SickKids Acute Care Transport Services team, for recruiting study subjects, Ashley Coppin for clinical data collection, the neonatal neurodevelopmental follow-up clinic and the EEG technicians from SickKids. This work was supported by the Canadian Institutes for Health Research [MOP-133710, PJT-166076], the National Institutes of Health [R01 HD101419], and the SickKids Foundation awarded to EWYT. SM is supported by the EU under the H2020 Research and Innovation Programme (Marie Skłodowska-Curie grant agreement, INFANS Project 813483). SV is supported by the Academy of Finland (335788, 332017, and 343498), Sigrid Juselius Foundation and Pediatric Foundation (Finland). ML is supported by the University of British Columbia Clinician Investigator Program.

Conflict of Interest

All authors declare no competing interests in direct relevance regarding the work related to this manuscript.

Author Contributions

ML, SM, EWYT, and SV contributed to conceptualization and study design. DK, EM, LL, CH, VC, and EWYT contributed to data collection. All authors contributed to data interpretation. SM and SV processed the EEG information and computed the BSN trends and outputs. SM performed the ROC, NPV, and PPV graphs. ML, SM, EWYT, and SV contributed to the statistical analysis, *ad hoc* visual analysis, and the creation of figures and tables, and they wrote the original draft. All authors contributed to the critical approval of the paper, including reviewing, and editing the final manuscript.

Data Availability Statement

All data-sharing requests regarding the cohort used should be submitted to EWYT for consideration. After publication, access to anonymized data might be granted for non-commercial research at the discretion of EWYT. Additional information related to building and validation of BSN is available on request and will be made available to interested research partners on reasonable request to SV; the prerequisite for data sharing is a data transfer agreement, approved by the legal departments of the requesting researcher and by all legal departments of the institutions that provided data for the study, and ethics clearance. The BSN solution is available via our cloud service (https://babacloud.fi/). The cloud interface needs credentials that are available at request from the SV. The system will not store the EEG files, and the user is encouraged to use only pseudonymized files for maximal data protection.

References

1. Kurinczuk JJ, White-Koning M, Badawi N. Epidemiology of neonatal encephalopathy and hypoxic–ischaemic encephalopathy. Early Hum Dev. 2010;86(6):329-338.
2. Chalak L, Hellstrom-Westas L, Bonifacio S, et al. Bedside and laboratory neuromonitoring in neonatal encephalopathy. Semin Fetal Neonatal Med. 2021;26(5):101273.
3. Steiner M, Urlesberger B, Giordano V, et al. Outcome prediction in neonatal hypoxic-ischaemic encephalopathy using neurophysiology and neuroimaging. Neonatology. 2022;119(4):483-493.
4. Goswami I, Guillot M, Tam EWY. Predictors of long-term neurodevelopmental outcome of hypoxic-ischemic encephalopathy treated with therapeutic hypothermia. Semin Neurol. 2020;40(3):322-334.
5. El-Dib M, Abend NS, Austin T, et al. Neuromonitoring in neonatal critical care part I: neonatal encephalopathy and neonates with possible seizures. Pediatr Res. 2022;94:64-73.
6. Awal MA, Lai MM, Azemi G, Boashash B, Colditz PB. EEG background features that predict outcome in term neonates with hypoxic ischaemic encephalopathy: a structured review. Clin Neurophysiol. 2016;127(1):285-296.
7. Han Y, Fu N, Chen W, et al. Prognostic value of electroencephalography in hypothermia-treated neonates with hypoxic-ischemic encephalopathy: a meta-analysis. Pediatr Neurol. 2019;93:3-10.
8. 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.
9. Dilena R, Raviglione F, Cantalupo G, et al. Consensus protocol for EEG and amplitude-integrated EEG assessment and monitoring in neonates. Clin Neurophysiol. 2021;132(4):886-903.
10. Ouwehand S, Smidt Lisanne CA, Dudink J, et al. Predictors of outcomes in hypoxic-ischemic encephalopathy following hypothermia: a meta-analysis. Neonatology. 2020;117(4):411-427.
11. Montazeri S, Airaksinen M, Nevalainen P, et al. An automated bedside measure for monitoring neonatal cortical activity: a supervised deep learning-based electroencephalogram classifier with external cohort validation. Lancet Digital Health. 2022;4(12):e884-e892.
12. Montazeri S, Nevalainen P, Metsäranta M, Stevenson NJ, Vanhatalo S. Clinical outcome prediction with an automated EEG trend, brain state of the newborn, after perinatal asphyxia. Clin Neurophysiol. 2024;162:68-76.
13. Kota S, Kang S, Liu Y-L, et al. Prognostic value of quantitative EEG in early hours of life for neonatal

encephalopathy and neurodevelopmental outcomes. Pediatr Res. 2024;96(3):685-694.
14. Kamino D, Widjaja E, Brant R, et al. Severity and duration of dysglycemia and brain injury among patients with neonatal encephalopathy. EClinicalMedicine. 2023;58:101914.
15. Badawi N, Kurinczuk JJ, Keogh JM, et al. Antepartum risk factors for newborn encephalopathy: the Western Australian case-control study. BMJ. 1998;317(7172):1549-1553.
16. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform. 2009;42(2):377-381.
17. Montazeri S, Pinchefsky E, Tse I, et al. Building an open source classifier for the neonatal EEG background: a systematic feature-based approach from expert scoring to clinical visualization. Front Hum Neurosci. 2021;15:675154.
18. Alharbi HM, Pinchefsky EF, Tran M-A, et al. Seizure burden and neurologic outcomes after neonatal encephalopathy. Neurology. 2023;100:e1976-e1984.
19. Webb L, Kauppila M, Roberts JA, Vanhatalo S, Stevenson NJ. Automated detection of artefacts in neonatal EEG with residual neural networks. Comput Methods Prog Biomed. 2021;208:106194.
20. Dilena R, Cilio MR. Free access via computational cloud to deep learning-based EEG assessment in neonatal hypoxic-ischemic encephalopathy: revolutionary opportunities to overcome health disparities. Pediatr Res. 2024;96:841-843.
21. Arbelaez Ossa L, Starke G, Lorenzini G, Vogt JE, Shaw DM, Elger BS. Re-Focusing Explainability in Medicine. SAGE Publications; 2022.
22. Weeke LC, Groenendaal F, Mudigonda K, et al. A novel magnetic resonance imaging score predicts neurodevelopmental outcome after perinatal asphyxia and therapeutic hypothermia. J Pediatr. 2018;192:33-40.e2.
23. Del Rosario C, Slevin M, Molloy EJ, Quigley J, Nixon E. How to use the Bayley scales of infant and toddler development. Arch Dis Child Educ Pract Ed. 2020;106(2):108-112.
24. David WH Jr, Lemeshow S, Sturdivant RX. Applied Logistic Regression. 3rd ed. Wiley; 2013.
25. Oehmichen M, Woetzel F, Meissner C. Hypoxic-ischemic changes in SIDS brains as demonstrated by a reduction in MAP2-reactive neurons. Acta Neuropathol. 2009;117(3):267-274.
26. Del Bigio MR, Becker LE. Microglial aggregation in the dentate gyrus: a marker of mild hypoxic-ischaemic brain insult in human infants. Neuropathol Appl Neurobiol. 1994;20(2):144-151.
27. Chandrasekaran M, Chaban B, Montaldo P, Thayyil S. Predictive value of amplitude-integrated EEG (aEEG) after rescue hypothermic neuroprotection for hypoxic ischemic encephalopathy: a meta-analysis. J Perinatol. 2017;37(6):684-689.
28. Basti C, Maranella E, Cimini N, et al. Seizure burden and neurodevelopmental outcome in newborns with hypoxic-ischemic encephalopathy treated with therapeutic hypothermia: a single center observational study. Seizure. 2020;83:154-159.
29. Obeid R, Tsuchida TN. Treatment effects on neonatal EEG. J Clin Neurophysiol. 2016;33(5):376-381.
30. Liu W, Yang Q, Wei H, Dong W, Fan Y, Hua Z. Prognostic value of clinical tests in neonates with hypoxic-ischemic encephalopathy treated with therapeutic hypothermia: a systematic review and meta-analysis. Front Neurol. 2020;11:133.
31. Morales MM, Montaldo P, Ivain P, et al. Association of Total Sarnat Score with brain injury and neurodevelopmental outcomes after neonatal encephalopathy. Arch Dis Child Fetal Neonatal Ed. 2021;106(6):669-672.
32. Cainelli E, Trevisanuto D, Cavallin F, Manara R, Suppiej A. Evoked potentials predict psychomotor development in neonates with normal MRI after hypothermia for hypoxic-ischemic encephalopathy. Clin Neurophysiol. 2018;129(6):1300-1306.
33. Wu YW, Monsell SE, Glass HC, et al. How well does neonatal neuroimaging correlate with neurodevelopmental outcomes in infants with hypoxic-ischemic encephalopathy? Pediatr Res. 2023;94(3):1018-1025.
34. Marlow N. Measuring neurodevelopmental outcome in neonatal trials: a continuing and increasing challenge. Arch Dis Child Fetal Neonatal Ed. 2013;98(6):F554-F558.
35. Kilbride HWMD, Aylward GPPA, Carter BMD. What are we measuring as outcome? Looking beyond neurodevelopmental impairment. Clin Perinatol. 2018;45(3):467-484.
36. Merz EC, Wiltshire CA, Noble KG. Socioeconomic inequality and the developing brain: spotlight on language and executive function. Child Dev Perspect. 2019;13(1):15-20.
37. Benavente-Fernández I, Siddiqi A, Miller SP. Socioeconomic status and brain injury in children born preterm: modifying neurodevelopmental outcome. Pediatr Res. 2020;87(2):391-398.
38. Ilardi D, Sanz JH, Cassidy AR, et al. Neurodevelopmental evaluation for school-age children with congenital heart disease: recommendations from the cardiac neurodevelopmental outcome collaborative. Cardiol Young. 2020;30(11):1623-1636.
39. Hammer CS, Komaroff E, Rodriguez BL, Lopez LM, Scarpino SE, Goldstein B. Predicting Spanish-English bilingual children's language abilities. J Speech Lang Hear Res. 2012;55(5):1251-1264.

40. Zablotsky B, Black LI, Maenner MJ, et al. Prevalence and trends of developmental disabilities among children in the United States: 2009–2017. Pediatrics. 2019;144(4):1-11.
41. Bitta M, Kariuki SM, Abubakar A, Newton CRJC. Burden of neurodevelopmental disorders in low and middle-income countries: a systematic review and meta-analysis. Wellcome Open Research. 2018;2:121.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix S1.

