

## RESEARCH ARTICLE

# Machine learning for the early prediction of infants with electrographic seizures in neonatal hypoxic-ischemic encephalopathy

Andreea M. Pavel<sup>1,2</sup>  | John M. O'Toole<sup>1,2</sup>  | Jacopo Proietti<sup>1</sup> | Vicki Livingstone<sup>1,2</sup> | Subhabrata Mitra<sup>3</sup> | William P. Marnane<sup>1,4</sup> | Mikael Finder<sup>5,6</sup>  | Eugene M. Dempsey<sup>1,2</sup> | Deirdre M. Murray<sup>1,2</sup> | Geraldine B. Boylan<sup>1,2</sup>  | ANSeR Consortium\*

<sup>1</sup>INFANT Research Centre, University College Cork, Cork, Ireland

<sup>2</sup>Department of Paediatrics and Child Health, University College Cork, Cork, Ireland

<sup>3</sup>Institute for Women's Health, University College London, London, UK

<sup>4</sup>Electrical & Electronic Engineering, School of Engineering, University College Cork, Cork, Ireland

<sup>5</sup>Department of Neonatal Medicine, Karolinska University Hospital, Stockholm, Sweden

<sup>6</sup>Division of Paediatrics, Department CLINTEC, Karolinska Institutet, Stockholm, Sweden

## Correspondence

Geraldine B. Boylan, Department of Paediatrics and Child Health, INFANT Research Centre, University College Cork, Paediatric Academic Unit, Cork University Hospital, 2nd floor, Wilton, Cork, Ireland.

Email: [g.boylan@ucc.ie](mailto:g.boylan@ucc.ie)

## Funding information

Health Research Board, Grant/Award Number: NEPTuNE CDA-2018-008; Wellcome Trust, Grant/Award Number: 098983 and 209325

## Abstract

**Objective:** To assess if early clinical and electroencephalography (EEG) features predict later seizure development in infants with hypoxic-ischemic encephalopathy (HIE).

**Methods:** Clinical and EEG parameters <12h of birth from infants with HIE across eight European Neonatal Units were used to develop seizure-prediction models. Clinical parameters included intrapartum complications, fetal distress, gestational age, delivery mode, gender, birth weight, Apgar scores, assisted ventilation, cord pH, and blood gases. The earliest EEG hour provided a qualitative analysis (discontinuity, amplitude, asymmetry/asynchrony, sleep-wake cycle [SWC]) and a quantitative analysis (power, discontinuity, spectral distribution, inter-hemispheric connectivity) from full montage and two-channel amplitude-integrated EEG (aEEG). Subgroup analysis, only including infants without anti-seizure medication (ASM) prior to EEG was also performed. Machine-learning (ML) models (random forest and gradient boosting algorithms) were developed to predict infants who would later develop seizures and assessed using Matthews correlation coefficient (MCC) and area under the receiver-operating characteristic curve (AUC).

**Results:** The study included 162 infants with HIE (53 had seizures). Low Apgar, need for ventilation, high lactate, low base excess, absent SWC, low EEG power, and increased EEG discontinuity were associated with seizures. The following predictive models were developed: clinical (MCC 0.368, AUC 0.681), qualitative EEG (MCC 0.467, AUC 0.729), quantitative EEG (MCC 0.473, AUC 0.730), clinical and qualitative EEG (MCC 0.470, AUC 0.721), and clinical and quantitative EEG (MCC 0.513, AUC 0.746). The clinical and qualitative-EEG model significantly outperformed the clinical model alone (MCC 0.470 vs 0.368, *p*-value

\*Membership of the ANSeR Consortium is provided in the Acknowledgments.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Epilepsia* published by Wiley Periodicals LLC on behalf of International League Against Epilepsy.

.037). The clinical and quantitative-EEG model significantly outperformed the clinical model (MCC 0.513 vs 0.368,  $p$ -value .012). The clinical and quantitative-EEG model for infants without ASM ( $n = 131$ ) had MCC 0.588, AUC 0.832. Performance for quantitative aEEG ( $n = 159$ ) was MCC 0.381, AUC 0.696 and clinical and quantitative aEEG was MCC 0.384, AUC 0.720.

**Significance:** Early EEG background analysis combined with readily available clinical data helped predict infants who were at highest risk of seizures, hours before they occur. Automated quantitative-EEG analysis was as good as expert analysis for predicting seizures, supporting the use of automated assessment tools for early evaluation of HIE.

#### KEYWORDS

machine learning, neonatal encephalopathy, neonatal seizures, neonates, prediction algorithm

## 1 | INTRODUCTION

Neonatal encephalopathy caused by perinatal hypoxia-ischemia is a worldwide health problem and a major cause of mortality and morbidity, with incidence ranging from 1.5–3 per 1000 live births in high-income countries to 14.9 per 1000 live births in low-income countries.<sup>1,2</sup> Brain injury following hypoxic-ischemic injury evolves over subsequent hours and days and is associated with changing clinical and electroencephalography (EEG) features.<sup>3,4</sup> Hypoxic-ischemic encephalopathy (HIE) is the main cause of seizures in full-term infants and more than half of those with moderate to severe encephalopathy develop seizures.<sup>5,6</sup> The only treatment currently recommended for moderate and severe HIE is therapeutic hypothermia initiated within 6 h of birth.<sup>7–9</sup> This intervention has been shown to improve long-term neurodevelopmental outcomes and decrease total seizure burden (TSB).<sup>10,11</sup> However, even after the introduction of hypothermia, the incidence of seizures in neonates with HIE was estimated by several studies to be between 30% and 65%.<sup>10,12,13</sup> One study reported that high seizure burden in infants with HIE was associated with poor neurodevelopmental outcome, independent of the severity of encephalopathy.<sup>14</sup> In addition, regardless of the background etiology, the presence of seizures in neonates is associated with worse neurodevelopmental outcomes and increased mortality.<sup>15,16</sup>

To improve outcomes in this population, prompt recognition and treatment of seizures is vital.<sup>17,18</sup> However the clinical diagnosis of seizures in neonates is challenging, due to the high frequency of electrographic-only seizures in HIE and the “uncoupling” phenomenon following anti-seizure medication (ASM).<sup>19–21</sup> Even when clinical manifestations are present, it can be difficult to differentiate seizures from normal neonatal behaviors.<sup>21</sup> The gold standard for seizure diagnosis is continuous video-EEG

#### Key points

- Early electroencephalography (EEG) background combined with clinical information can predict infants which later develop seizure in hypoxic-ischemic encephalopathy (HIE).
- Automated quantitative-EEG analysis was as predictive as an expert neurophysiologist analysis for predicting seizures, supporting the use of automated assessment tools.
- The current machine-learning (ML) prediction models could be used as objective bedside tools to identify which neonates with HIE are at the highest risk of seizures.

monitoring, as recommended by the American Clinical Neurophysiology Society (ACNS) and the International League Against Epilepsy (ILAE).<sup>22,23</sup> Unfortunately, continuous EEG monitoring requires specialized equipment and expert personnel, and is used only for short recordings in most neonatal centers, thereby reducing the likelihood of seizure detection. To overcome this and improve the allocation of scarce resources, a solution would be early identification of infants who are vulnerable to develop seizures, which should also reduce the delays in treatment and lead to better outcomes.<sup>24</sup>

Previous attempts were made to develop seizure prediction models using clinical, biochemical, and EEG markers, alone or in combination.<sup>25–33</sup> A combination of Apgar score at 5 min, pH < 7.0 and delivery room intubation,<sup>25,27</sup> and EEG background analysis with or without other clinical parameters<sup>28–32</sup> was previously investigated. All previous attempts were hampered by small, heterogenous neonatal populations, or by lack

of prospective, detailed, and early EEG monitoring. Traditionally, logistic regression has been used to develop predictive models for newborns with seizures.<sup>30,31</sup> However, in recent years, machine learning (ML) techniques have allowed for a more robust analysis of complex data, resulting in development of decision support tools for healthcare professionals.<sup>33–36</sup>

The aim of this study was to use ML techniques to assess the ability of early clinical parameters and EEG background features, to predict those infants with HIE who later develop seizures, using a large multicenter data set. We hypothesized that an ML model combining early clinical data and EEG background features would have the best ability to predict the infants with HIE at risk of seizures.

## 2 | MATERIAL AND METHODS

### 2.1 | Study setting and participants

This is a secondary data analysis of infants recruited for two prospective, multicenter, cohort studies ([ClinicalTrials.gov](https://clinicaltrials.gov) Identifier: NCT02160171 and NCT02431780) from eight tertiary-level neonatal intensive care units across Europe (Ireland, United Kingdom, Sweden, The Netherlands).<sup>6,36</sup> Ethics committee approval from each recruiting site was obtained before recruitment commenced. Both studies included infants born at  $\geq 36$  weeks of gestation, requiring continuous conventional EEG (cEEG) monitoring because they were at high risk of developing seizures. For this analysis we included infants with a clinical and electrographic diagnosis of HIE who had cEEG recording within 12 h of birth and before the development of any electrographic seizures. The diagnosis of HIE was based on an antenatal history of a hypoxic-ischemic event and the presence of encephalopathy on early neurologic examination (most severe modified Sarnat score within 24 h of life), corroborated with EEG background and/or magnetic resonance imaging (MRI) changes consistent with HIE injury. Infants with encephalopathy from other causes, such as sepsis, meningitis, stroke, metabolic, or genetic, or a combined diagnosis, were excluded from the analysis. In the primary studies, an infant was considered to have seizures if experts identified at least one seizure lasting at least 10 s throughout the EEG monitoring. For this study, infants with only very brief seizures ( $< 30$  s) and a total seizure burden (TSB)  $< 2$  min throughout the EEG monitoring were excluded from the analysis on the basis that generally these seizures would not have been treated. A subgroup analysis was also performed excluding all infants who received ASM before the EEG because of the known effect of ASMs on the background EEG pattern.<sup>37,38</sup>

### 2.2 | Early clinical features

The following clinical parameters, available within 6 h of birth, were included in the prediction models: intrapartum complications, suspected fetal distress in labor, gestational age, mode of delivery, gender, birth weight, Apgar scores at 1, 5, and 10 min, assisted ventilation at 10 min of age, lowest cord pH, first postnatal lactate, and first postnatal base excess. Any of the following were considered as an intrapartum complication: placental abruption, ruptured uterus, vasa praevia, intrapartum hemorrhage, cord accident or prolapse, shoulder dystocia, meconium-stained liquor, or other (poor progression, HELLP syndrome [hemolysis, elevated liver enzymes, low platelet count]/eclampsia, complicated breech presentation, reduced fetal movements, fetal bradycardia, prolonged rupture of membranes).

### 2.3 | EEG monitoring and analysis

All infants had cEEG monitoring commencing as soon as possible after birth using NicoletOne ICU Monitor (Natus, USA), Nihon Kohden EEG (Neurofax EEG-1200, Japan), or XLTek EEG (Natus, USA). Personnel were trained at each study site to position disposable electrodes at F3, F4, C3, C4, Cz, T3, T4, O1/P3, and O2/P4, according to 10:20 EEG electrode neonatal system. The EEG recording had a sampling rate of 250 Hz or 256 Hz, with a filter bandwidth between 0.5 and 70 Hz for review.

For each infant, we extracted the earliest 1-h epoch of good-quality EEG recording available before 12 h of age and at least 1 h before the onset of electrographic seizures. All epochs were reviewed, and artifacts were annotated and removed from the quantitative-EEG analysis. For each EEG epoch included in the analysis, we performed a qualitative (visual) and a quantitative analysis.

The *qualitative-EEG analysis* was performed by two expert neurophysiologists (GBB and JP) blinded to the seizure status of the infant. The following features were assessed individually as present or absent as per the ACNS<sup>39</sup>: 1, any abnormal discontinuity; 2, discontinuity  $> 10$  s; 3, continuous low voltage to isoelectric; 4, asymmetry and/or asynchrony; 5, sleep-wake cycles (SWCs) in the first 12 h. Abnormal discontinuity was defined as persistent intervals (interburst intervals [IBIs]) of relatively lower amplitude lasting more than 6 s. Epochs with more marked discontinuity, with  $IBI > 10$  s were analyzed separately. A cutoff of  $25 \mu\text{V}$  or less was used to define low voltage, and EEG activity below  $2.5 \mu\text{V}$  in amplitude was referred to as isoelectric activity. Asymmetry was defined as disparity in voltage (more than 2:1) or background feature distribution between homologous areas of the two hemispheres. Asynchrony was defined

as noncoherent occurrence of EEG activities ( $\geq 1.5$  s difference in onset of bursts) over regions on the same or opposite sides of the head. Any differences between neurophysiologists were discussed and a consensus was reached.

Electrographic seizures were defined according to Clancy et al. as at least one EEG channel with sudden, repetitive, and evolving waveforms for a minimum 10 s.<sup>40</sup> For all infants, electrographic seizures were identified by the same two neurophysiology experts, as described previously.<sup>6,36</sup> Seizure characteristics were calculated based on these annotations: seizure number, TSB (all seizures during the entire EEG monitoring period, minutes), maximum seizure burden (maximum seizure burden within an hour, minutes/hour), status epilepticus (seizure burden of  $>30$  min within 1 h), seizure period (hours from the beginning of first electrographic seizure to the end of last seizure).

*Quantitative-EEG analysis* was performed using the NEURAL software package (version 0.4.3), extracting a set of features described previously by our group: power from the 1-h cEEG epoch, discontinuity, spectral distribution, and inter-hemispheric connectivity features.<sup>41</sup> The power features included absolute spectral power and range EEG (median, upper, and lower margin). The discontinuity features included range-EEG asymmetry, IBI analysis (maximum and median length, percentage, and number of IBIs), amplitude skewness, and kurtosis. IBIs were detected using an algorithm developed for preterm infants<sup>42</sup> that was also validated in an HIE term cohort.<sup>43</sup> The spectral distribution features included spectral relative power, spectral flatness, spectral difference, spectral edge frequency, and fractal dimension. Inter-hemispheric connectivity analysis included connectivity brain symmetry index (BSI) and connectivity coherence. Spectral power (absolute and relative), amplitude skewness and kurtosis, spectral flatness and difference, and coherence and connectivity BSI were generated separately for four frequency bands: 1 to 4 Hz (delta), 4 to 7 Hz (theta), 7 to 13 Hz (alpha), and 13 to 30 Hz (beta). Range EEG is generated for a 1 to 20 Hz band and fractal dimension is generated for 1 to 30 Hz.

Quantitative analysis was also performed using two-channel EEG from F3-C3 and F4-C4. Due to increased usability of two-channel amplitude integrated amplitude EEG systems (aEEG) in neonatal units, we have developed predictive models separately using quantitative analysis from a reduced montage (F3-C3 and F4-C4).

## 2.4 | Statistical analysis

Categorical variables were described using frequencies and percentages and continuous variables using means and standard deviations (SDs) when the variables were

normally distributed, or medians and interquartile ranges (IQRs) otherwise. For comparisons of continuous variables between two groups (seizure vs nonseizure; moderate HIE vs severe HIE) the independent-sample *t* test was used for normally distributed variables and the Mann-Whitney *U* test for non-normally distributed variables. The chi-square test or Fisher's exact test (in the case of small, expected counts) was used for categorical variables. All tests were two sided and a *p*-value  $< .05$  was considered statistically significant. IBM SPSS Statistics (version 24.0, IBM Corp., Armonk, NY, USA) was used for the statistical analysis.

## 2.5 | Machine-learning (ML) analysis

ML models were developed (using Python 3.10.6) separately and in combination for clinical features, and qualitative EEG and quantitative EEG features, to predict infants who later developed seizures. The models were developed using either bagging or boosting ensembles of decision trees. Random forests (a bagging approach) were used to develop models for clinical and qualitative EEG features. Gradient boosting was used for quantitative EEG features. These models were trained and tested using a leave-one-out cross-validation procedure. The depth of the decision tree was optimized from the data. This controlled a level of regularization for the models: A deeper tree captures more information from the feature set but may cause the model to over-fit. The parameter was selected from a grid search within a nested cross-validation. This nested scheme employed a 10-fold cross-validation repeated 5 times, with each repetition using a different 10-fold random selection.

### 2.5.1 | Clinical and qualitative EEG features

For the missing clinical data included in the ML model (suspected fetal distress, Apgar scores, assisted ventilation, cord pH, postnatal lactate, and base excess), the mean value of each feature was used for imputation of missing values. Features with more than 50% missing data were excluded. For both clinical and qualitative models, the default number of iterations was used to grow 100 decision trees. The maximum depth of the decision tree was selected in the nested cross-validation from the set {1,2,3,4}.

### 2.5.2 | Quantitative EEG features

To compensate for the high levels of discontinuous activity in some EEG records, we generated an extra feature

set weighted by the quantity of discontinuous activity. This extra feature set consisted of the existing features (excluding the IBI features) multiplied by the percentage of IBIs. We then combined both feature sets for use in building the ML model. ASM status (yes/no if administered before the start of the EEG epoch) and hypothermia status (yes/no if at time of EEG epoch) were also included as adjusting features in the quantitative model. In addition, separate ML models were developed on a subgroup of infants who did not receive any ASM prior to the EEG epoch used in the analysis. We selected the CatBoost implementation of gradient boosting to better manage over-fitting that can occur with boosting algorithms.<sup>44</sup> Because this is a small data set for ML (<200 data points) and model variance within the leave-one-out strategy can lead to poor performance, we reduced the number of iterations (number of trees) during training from 1000 (default value) to 40 and set the learning rate to 0.1. The depth of the decision tree was selected in the nested cross-validation from the set {2,3,...,6}.

### 2.5.3 | Assessment and comparison of ML models

To account for the imbalanced classes of seizure and nonseizure, Matthews correlation coefficient (MCC) was used to assess and compare performance of the ML models.<sup>45</sup> William's test for dependent correlations was used to compare MCCs. More standard metrics were also included: area under the receiver-operating characteristic curve (AUC), sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). Testing probabilities for the clinical and qualitative EEG models, clinical and quantitative EEG models, and clinical and quantitative aEEG models were then combined using a logit aggregation formula.<sup>46</sup> For the combined models (clinical and EEG/aEEG models) we used a "late-stage fusion" method. Agreement between the individual prediction models (clinical, qualitative EEG, quantitative EEG) to quantify for the potential overlap between the models was assessed using Cohen's kappa (values  $\leq 0$  indicates no agreement, 0.01–0.20 none to slight, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 substantial, and 0.81–1.00 almost perfect agreement).

## 3 | RESULTS

From 504 infants recruited in both studies, 266 infants had a diagnosis of HIE, and of those 164 infants had at least 1 h of EEG monitoring before 12 h of age and before emergence of electrographic seizures. Two infants had

individual seizures <30 s and a TSB <2 min and were excluded. Hence, 162 neonates with HIE of all grades—mild 62 (38.3%), moderate 69 (42.6%), and severe 31 (19.1%)—were included in the study analysis. Electrographic seizures were present in 53 infants (32.7%) (seizure group). Demographic and clinical characteristics overall and by seizure status group are presented in Table 1. The median (IQR) age at start of the EEG monitoring was 4.7 (3.3–7.7) hours and the median (IQR) age at the time the EEG epoch was analyzed was 6.4 (4.2–8.7) hours, and there were no significant differences in these measures between the seizure and nonseizure groups. Infants with seizures had lower Apgar scores at 1, 5, and 10 min; higher post-natal lactate; and lower base excess; and they were more likely to be cooled and to require assisted ventilation at 10 min of age compared with infants in the nonseizure group. Overall, 31 infants received ASM for clinical concerns of seizures prior to the EEG epoch analyzed.

Electrographic seizure characteristics are presented in Table 2. The median (IQR) age at first electrographic seizure was 14.4 (10.3–19.6) hours. Compared to infants with moderate HIE, infants with severe HIE had a significantly higher number of seizures (median (IQR) 39 (28–104) vs 11<sup>5–25</sup> seizures,  $p$ -value <.001), individual seizures were of shorter median duration (94 (68–140) vs 163 (93–838) seconds,  $p$ -value .012), TSB was higher (median (IQR) 126 (71–212) vs 59 (26–76) minutes,  $p$ -value .003), seizure period was longer (median (IQR) 40.2 (25.3–69.3) vs 9.8 (3.2–32.4) hours,  $p$ -value <.001), and they received more ASM (median (IQR) 3<sup>3,4</sup> vs 1<sup>1,2</sup> doses,  $p$ -value <.001). Eighteen infants (34%) had status epilepticus, and there were no significant differences between the moderate and severe HIE groups.

EEG features overall and by seizure status group are presented in Table 3 (see Table S1 for all EEG features included in ML models). Low voltage to isoelectric pattern on EEG and absence of SWC were associated with seizure emergence. Several quantitative-EEG features were significant predictors of seizures. Compared to infants from the non-seizure group, infants with seizures had significantly lower power (spectral power, for each band and range EEG), lower amplitude kurtosis in the beta band, lower relative power in delta and theta bands, and lower spectral difference in the delta and theta bands. The following features were significantly higher for infants with seizures compared to infants without seizures: range EEG asymmetry, IBI length and percentage, amplitude skewness in delta and theta bands, amplitude kurtosis in the delta band, spectral relative power in alpha and beta band, spectral flatness in beta, spectral difference in the beta band, spectral edge frequency and fractal dimension, connectivity BSI in delta and alpha bands, and mean connectivity coherence in the theta band.

**TABLE 1** Study sample demographics and clinical characteristics

|   | <i>n</i> | All infants   | <i>n</i> | Non-seizure group<br><i>n</i> = 109 | <i>n</i> | Seizure group<br><i>n</i> = 53 | <i>p</i> -value    |
|---|----------|---------------|----------|-------------------------------------|----------|--------------------------------|--------------------|
| Gestational age at birth (weeks), mean (SD) <sup>a</sup>        | 162      | 40.01 (1.32)  | 109      | 39.93 (1.29)                        | 53       | 40.18 (1.37)                   | .268 <sup>b</sup>  |
| Intrapartum complications (yes), <i>n</i> (%) <sup>a</sup>      | 162      | 140 (86.4)    | 109      | 91 (83.5)                           | 53       | 49 (92.5)                      | .118 <sup>d</sup>  |
| Suspected fetal distress (yes), <i>n</i> (%) <sup>a</sup>       | 145      | 107 (66)      | 102      | 76 (74.5)                           | 43       | 31 (72.1)                      | .762 <sup>d</sup>  |
| Mode of delivery, <i>n</i> (%) <sup>a</sup>                     |          |               |          |                                     |          |                                |                    |
| Unassisted vaginal delivery                                     | 162      | 46 (28.4)     | 109      | 30 (27.5)                           | 53       | 16 (30.2)                      | .962 <sup>d</sup>  |
| Assisted vaginal delivery                                       |          | 62 (38.3)     |          | 42 (38.5)                           |          | 20 (37.7)                      |                    |
| Elective cesarean section                                       |          | 8 (4.9)       |          | 5 (4.6)                             |          | 3 (5.7)                        |                    |
| Emergency cesarean section                                      |          | 46 (28.4)     |          | 32 (29.4)                           |          | 14 (26.4)                      |                    |
| Birth weight (g), mean (SD) <sup>a</sup>                        | 162      | 3508 (606)    | 109      | 3508 (608)                          | 53       | 3507 (609)                     | .991 <sup>b</sup>  |
| Male gender, <i>n</i> (%) <sup>a</sup>                          | 162      | 103 (63.6)    | 109      | 73 (67)                             | 53       | 30 (56.6)                      | .198 <sup>d</sup>  |
| Apgar at 1 min, median (IQR) <sup>a</sup>                       | 157      | 2 (1–3)       | 106      | 2 (1–3)                             | 51       | 1 (0–2)                        | <.001 <sup>c</sup> |
| Apgar at 5 min, median (IQR) <sup>a</sup>                       | 157      | 4 (2–5)       | 106      | 4 (3–6)                             | 51       | 3 (1–4)                        | <.001 <sup>c</sup> |
| Apgar at 10 min, median (IQR) <sup>a</sup>                      | 144      | 5 (4–7)       | 98       | 6 (5–8)                             | 46       | 4 (2–5)                        | <.001 <sup>c</sup> |
| Assisted ventilation at 10 min (yes), <i>n</i> (%) <sup>a</sup> | 160      | 109 (67.3)    | 108      | 65 (59.6)                           | 52       | 44 (83)                        | .005               |
| Lowest cord pH, mean (SD) <sup>a</sup>                          | 137      | 7.00 (0.2)    | 94       | 7.01 (0.19)                         | 43       | 6.99 (0.2)                     | .694 <sup>b</sup>  |
| First postnatal base excess, mean (SD) <sup>a</sup>             | 121      | −15.4 (5.8)   | 85       | −14.8 (5.3)                         | 36       | −17.1 (6.4)                    | .037 <sup>b</sup>  |
| First postnatal lactate, mean (SD) <sup>a</sup>                 | 102      | 11.8 (4.32)   | 69       | 10.9 (3.9)                          | 33       | 13.5 (4.7)                     | .003 <sup>b</sup>  |
| HIE grade at discharge  |          |               |          |                                     |          |                                |                    |
| Mild, <i>n</i> (%)  | 162      | 62 (38.3)     | 109      | 62 (56.9)                           | 53       | 0                              | <.001 <sup>d</sup> |
| Moderate, <i>n</i> (%)  |          | 69 (42.6)     |          | 39 (35.8)                           |          | 30 (56.6)                      |                    |
| Severe, <i>n</i> (%)  |          | 31 (19.1)     |          | 8 (7.3)                             |          | 23 (43.4)                      |                    |
| Therapeutic hypothermia (yes), <i>n</i> (%)                     | 162      | 133 (82.1)    | 109      | 81 (74.3)                           | 53       | 52 (98.1)                      | <.001 <sup>d</sup> |
| Age at start of therapeutic hypothermia (h), median (IQR)       | 133      | 2 (1–4)       | 81       | 2 (1–5)                             | 52       | 1.5 (1–3)                      | .070 <sup>c</sup>  |
| Age at start of EEG monitoring (h), median (IQR)                | 162      | 4.7 (3.3–7.7) | 109      | 5.1 (3.4–7.7)                       | 53       | 4.2 (2.9–7.1)                  | .282 <sup>c</sup>  |
| Age at start of EEG epoch (h), median (IQR)                     | 162      | 6.4 (4.2–8.7) | 109      | 6.6 (4.2–8.6)                       | 53       | 6.1 (4.2–9.5)                  | .877 <sup>c</sup>  |
| Any AEDs given before EEG epoch analyzed (yes), <i>n</i> (%)    | 162      | 31 (19.1)     | 109      | 17 (15.6)                           | 53       | 14 (26.4)                      | .101 <sup>d</sup>  |

Note: *p*-value <.05 was considered statistically significant.

Abbreviations: AED, anti-epileptic drugs; EEG, electroencephalography; HIE, hypoxic-ischemic encephalopathy; ML, machine learning.

<sup>a</sup>Variables included in ML model.

<sup>b</sup>*p*-Value from independent sample *t* test for parametric data.

<sup>c</sup>*p*-Value from Mann-Whitney test for nonparametric data.

<sup>d</sup>*p*-Value from chi-square test or Fisher's exact test for categorical data.

### 3.1 | Machine learning (ML) models

Performance of the ML models are presented in Table 4. The MCC (95% CI) for the clinical model was 0.368 (0.219 to 0.506), 0.467 (0.319 to 0.611) for the qualitative-EEG model, and 0.473 (0.337 to 0.612) for the quantitative-EEG model. No significance difference was found between the clinical and EEG models: clinical vs qualitative-EEG models (*p*-value .083), clinical vs quantitative-EEG models (*p*-value .074), and qualitative-EEG vs quantitative-EEG

models (*p*-value .425). Cohen's kappa between the clinical model and the quantitative-EEG model was 0.47 and between the clinical model and qualitative-EEG model was 0.45. The EEG models (qualitative-EEG vs quantitative-EEG models) had a larger Cohen's kappa of 0.89, indicative of the similarity of the models in predicting seizures.

Combining the clinical models with the EEG models increased performance: MCC (95% CI) for the combined clinical and qualitative-EEG model is 0.470 (0.336 to 0.602) and 0.513 (0.376 to 0.645) for the clinical and EEG model. Both

**TABLE 2** Seizure characteristics by HIE severity ( $n = 53$ )

|   | All infants with seizures |                          |                        | p-value <sup>a</sup> |
|---|---------------------------|--------------------------|------------------------|----------------------|
|   | $n = 53$                  | Moderate HIE<br>$n = 30$ | Severe HIE<br>$n = 23$ |                      |
| Age at start of EEG monitoring (h), median (IQR)      | 4.1 (2.9–7.1)             | 4.3 (3.2–6.6)            | 4.1 (2.4–7.4)          | .566                 |
| EEG total monitoring (h), median (IQR)                | 94.5 (86.4–108.1)         | 93.6 (86.4–102.6)        | 95.9 (85.8–120.8)      | .44                  |
| Age at first electrographic seizure (h), median (IQR) | 14.4 (10.3–19.6)          | 13.5 (8.6–19.4)          | 14.6 (13.5–19.8)       | .229                 |
| Number of seizures, median (IQR)                      | 21 (8–51)                 | 11 (5–25)                | 39 (28–104)            | <.001                |
| Median seizure duration (s), median (IQR)             | 109 (79–355)              | 163 (93–838)             | 94 (68–140)            | .012                 |
| Total seizure burden (min), median (IQR)              | 71 (33–140)               | 59 (26–76)               | 126 (71–212)           | .003                 |
| Maximum seizure burden (min/h), median (IQR)          | 23 (15–32)                | 22 (16–37)               | 25 (13–32)             | .816                 |
| Status epilepticus, n (%)                             | 18 (34)                   | 9 (30%)                  | 9 (39.1)               | .487 <sup>b</sup>    |
| Seizure period (h), median (IQR)                      | 25.3 (8.4–42.7)           | 9.8 (3.2–32.4)           | 40.2 (25.3–69.3)       | <.001                |
| Received AED treatment at any time (yes), n (%)       | 49 (92.5)                 | 26 (86.7)                | 23 (100)               | .124 <sup>b</sup>    |
| AED before EEG monitoring, n (%)                      | 12 (22.6)                 | 7 (23.3)                 | 5 (21.7)               | .891 <sup>b</sup>    |
| Total AED doses after EEG start, median (IQR)         | 2 (1–4)                   | 1 (1–2)                  | 3 (3–4)                | <.001                |

Note: p-value <.05 was considered statistically significant.

Abbreviations: HIE, hypoxic-ischemic encephalopathy; EEG, electroencephalography; AED, anti-epileptic drugs; IQR, interquartile range.

<sup>a</sup>p-Value from Mann-Whitney test for nonparametric data unless otherwise stated.

<sup>b</sup>p-Value from chi-square test for categorical data.

models had statistically significant better performances than the clinical model alone: clinical and qualitative-EEG vs clinical (MCC 0.470 vs 0.368,  $p$ -value .037); clinical and quantitative-EEG vs clinical (MCC 0.513 vs 0.368,  $p$ -value .012). MCCs between clinical and qualitative-EEG vs quantitative-EEG models ( $p$ -value .475) and between clinical and quantitative-EEG vs quantitative-EEG models ( $p$ -value .100) did not differ significantly.

A calibration curve between the clinical model, the quantitative-EEG model, and the combined clinical and quantitative-EEG model is shown as Supplemental Material (Figure S1).

Performances for the ML models using the subgroup of infants with no ASM and aEEG monitoring are presented in Tables 5 and 6, respectively. The aEEG models included 159 infants (due to artifacts on the raw aEEG channels we could not produce an output from 3 infants) and no ASM subgroup included 131 infants (31 infants received at least one dose of ASM before the EEG epoch analyzed). The quantitative-aEEG model (MCC 0.381) and the clinical and quantitative-aEEG model (MCC 0.384) had similar performance. For the subgroup of infants without ASM, the clinical and qualitative-EEG model had an MCC of 0.588.

## 4 | DISCUSSION

In this large multicenter cohort of infants with HIE, 53 of 162 infants developed electrographic seizures at

a median age of 14.4 h. Using ML we investigated the ability of the early EEG (recorded at a median age of 6.4 h) and relevant clinical data to predict which neonates would later develop electrographic seizures. Several individual clinical parameters were associated with occurrence of seizures in HIE: low Apgar scores, need for resuscitation at birth, low base deficit, and high lactate. Visual EEG analysis revealed that low voltage and absence of SWC were predictive of seizures in addition to specific power, discontinuity, and spectral shape features on quantitative analysis. Of the individual ML models developed, the EEG models had similar performance (MCCs of 0.467 for qualitative and 0.473 for quantitative), and both outperformed the clinical model alone (MCC 0.368), although these differences were not statistically significant. However, the combined clinical and qualitative-EEG model significantly outperformed the quantitative-EEG model alone and the clinical and qualitative-EEG model significantly outperformed the clinical model alone. Of interest, although the aEEG model did not perform as well as conventional EEG models, performance improved from the clinical model alone, highlighting the value of any EEG monitoring. The combined clinical and quantitative-EEG model developed exclusively from infants with no prior ASM ( $n = 131$ ) had the best performance (MCC 0.588).

Several previous studies have investigated the predictive value of clinical and EEG parameters for seizure occurrence in neonates. Apgar scores, need for respiratory

**TABLE 3** Qualitative and quantitative-EEG features by seizure status

|  | Non-seizure group (n = 109) | Seizure group (n = 53) | p-value <sup>a</sup> |
|--|-----------------------------|------------------------|----------------------|
| <b>Qualitative-EEG analysis</b>        | <b>n (%)</b>                | <b>n (%)</b>           |                      |
| Discontinuity (yes)                    | 67 (61.5)                   | 26 (50.9)              | .134                 |
| Discontinuity >10 s (yes)              | 31 (28.4)                   | 22 (41.5)              | .096                 |
| Low voltage to isoelectric (yes)       | 22 (20.2)                   | 36 (67.9)              | <.001                |
| Asymmetry and/or asynchrony (yes)      | 5 (4.6)                     | 3 (5.7)                | .767                 |
| SWC in the first 12 h (yes)            | 16 (14.7)                   | 1 (1.9)                | .013                 |
| <b>Quantitative-EEG Features</b>       | <b>Median (IQR)</b>         | <b>Median (IQR)</b>    |                      |
| <i>Power features</i>                  |                             |                        |                      |
| Spectral power delta band              | 41.8 (19.3–66.9)            | 5.32 (2.18–29.6)       | <.001                |
| Spectral power theta band              | 8.01 (3.96–11.3)            | 0.934 (0.427–5.55)     | <.001                |
| Spectral power alpha band              | 3.29 (1.78–4.76)            | 0.648 (0.316–2.60)     | <.001                |
| Spectral power beta band               | 1.55 (0.95–2.97)            | 0.727 (0.253–1.75)     | <.001                |
| Range EEG lower margin                 | 15.5 (8.20–19.7)            | 5.61 (4.35–11.4)       | <.001                |
| Range EEG median                       | 28.8 (17.3–34.5)            | 8.94 (6.26–21.5)       | <.001                |
| Range EEG upper margin                 | 55.3 (45.2–68.0)            | 26.0 (14.1–48.7)       | <.001                |
| <i>Discontinuity features</i>          |                             |                        |                      |
| Range EEG asymmetry                    | 0.321 (0.277–0.443)         | 0.471 (0.295–0.644)    | .004                 |
| IBI length maximum                     | 4.84 (3.27–12.3)            | 50.9 (8.02–473.0)      | <.001                |
| IBI length median                      | 1.88 (1.59–4.12)            | 16.2 (2.66–48.0)       | <.001                |
| IBI percentage                         | 5.37 (1.03–51.6)            | 90.5 (33.0–98.1)       | <.001                |
| Amplitude skewness delta band          | 0.140 (0.120–0.183)         | 0.208 (0.135–0.313)    | .001                 |
| Amplitude skewness theta band          | 0.011 (0.009–0.017)         | 0.019 (0.011–0.031)    | <.001                |
| Amplitude kurtosis delta band          | 4.95 (4.44–6.82)            | 6.22 (4.49–12.1)       | .027                 |
| Amplitude kurtosis beta band           | 4.88 (4.04–5.98)            | 4.20 (3.34–5.03)       | .006                 |
| <i>Spectral shape features</i>         |                             |                        |                      |
| Spectral relative power delta band     | 74.9 (70.4–78.9)            | 69.5 (58.1–76.1)       | <.001                |
| Spectral relative power theta band     | 13.1 (11.0–15.5)            | 12.0 (9.80–14.0)       | .039                 |
| Spectral relative power alpha band     | 6.36 (5.18–7.65)            | 7.96 (5.58–10.5)       | .002                 |
| Spectral relative power beta band      | 3.44 (2.51–5.45)            | 8.44 (3.59–17.7)       | <.001                |
| Spectral flatness beta band            | 0.757 (0.674–0.814)         | 0.846 (0.766–0.884)    | <.001                |
| Spectral difference delta band         | 0.010 (0.008–0.013)         | 0.007 (0.003–0.013)    | .009                 |
| Spectral difference theta band         | 0.021 (0.013–0.025)         | 0.015 (0.006–0.022)    | .009                 |
| Spectral difference beta band          | 0.009 (0.006–0.012)         | 0.011 (0.007–0.015)    | .032                 |
| Spectral edge frequency                | 10.4 (8.75–13.0)            | 17.4 (10.0–22.7)       | <.001                |
| Fractal dimension                      | 1.44 (1.40–1.52)            | 1.65 (1.48–1.76)       | <.001                |
| <i>Connectivity features</i>           |                             |                        |                      |
| Connectivity BSI delta band            | 0.215 (0.173–0.272)         | 0.257 (0.216–0.324)    | <.001                |
| Connectivity BSI alpha band            | 0.218 (0.185–0.274)         | 0.254 (0.213–0.327)    | .017                 |
| Connectivity coherence mean theta band | 0.052 (0.040–0.089)         | 0.072 (0.045–0.154)    | .025                 |

Note: p-value <.05 was considered statistically significant. Frequency bands: delta: 0.5 to 3 Hz; theta: 4 to 7 Hz; alpha: 8 to 12 Hz; beta: 13 to 30 Hz.

Abbreviations: SWC, sleep-wake cycle; EEG, electroencephalography; IBI, interburst interval; BSI, brain symmetry index.

<sup>a</sup>p-Value from Mann-Whitney test for nonparametric data and chi-square test for categorical data.

support, and degree of metabolic acidosis were not found to be reliable predictors of seizures in neonates with HIE.<sup>25,27,28</sup> Our clinical model included antenatal and

postnatal clinical information available within 6 h of birth (the window to start therapeutic hypothermia). The performance of our clinical model (AUC 0.681) was very



**TABLE 4** Performance of machine-learning models ( $n = 162$ )

| ML model                            | MCC (95% CI)           | AUC (95% CI)           | Sensitivity | Specificity | PPV  | NPV  |
|-------------------------------------|------------------------|------------------------|-------------|-------------|------|------|
| Clinical model                      | 0.368 (0.219 to 0.506) | 0.681 (0.605 to 0.759) | 60.4        | 77.1        | 56.1 | 80.0 |
| Qualitative-EEG model               | 0.467 (0.319 to 0.611) | 0.729 (0.669 to 0.815) | 67.9        | 79.8        | 62.1 | 83.7 |
| Quantitative-EEG model              | 0.473 (0.337 to 0.612) | 0.730 (0.671 to 0.811) | 69.8        | 78.9        | 61.7 | 84.3 |
| Clinical and qualitative-EEG model  | 0.470 (0.336 to 0.602) | 0.721 (0.681 to 0.813) | 79.2        | 70.6        | 56.8 | 87.5 |
| Clinical and quantitative-EEG model | 0.513 (0.376 to 0.645) | 0.746 (0.700 to 0.833) | 75.5        | 78.0        | 62.5 | 86.7 |

Note: Statistical significance between models performance (MCC) using William's test for dependent correlations: clinical vs qualitative-EEG models,  $p$ -value .083; clinical vs quantitative-EEG models,  $p$ -value .074; qualitative-EEG vs quantitative-EEG models,  $p$ -value .425; clinical and qualitative-EEG vs qualitative-EEG models,  $p$ -value .475; clinical and qualitative-EEG vs clinical models,  $p$ -value .037; clinical and quantitative-EEG vs qualitative-EEG models,  $p$ -value .100; clinical and quantitative-EEG vs clinical models,  $p$ -value .012. Probability threshold was calculated from the receiver-operating characteristic curve as the point of equal sensitivity and specificity.

Abbreviations: ML, machine learning; MCC, Matthews correlation coefficient; AUC, area under the receiver-operating characteristic curve; PPV (NPV), positive (negative) predictive value.

**TABLE 5** Subgroup analysis of infants with no ASM given prior to the EEG epoch<sup>a</sup> ( $n = 131$ )

| ML model                            | MCC   | AUC   | Sensitivity | Specificity | PPV  | NPV  |
|-------------------------------------|-------|-------|-------------|-------------|------|------|
| Clinical model                      | 0.461 | 0.732 | 64.1        | 82.6        | 61   | 84.4 |
| Qualitative-EEG model               | 0.504 | 0.706 | 64.1        | 85.9        | 65.8 | 84.9 |
| Quantitative-EEG model              | 0.475 | 0.806 | 71.8        | 78.3        | 58.3 | 86.7 |
| Clinical and qualitative-EEG model  | 0.519 | 0.782 | 76.9        | 78.3        | 60   | 88.9 |
| Clinical and quantitative-EEG model | 0.588 | 0.832 | 66.7        | 90.2        | 74.3 | 86.5 |

Abbreviations: ML, machine learning; MCC, Matthews correlation coefficient; AUC, area under the receiver-operating characteristic curve; PPV (NPV), positive (negative) predictive value.

<sup>a</sup>Subgroup analysis of 131 infants (excluding 31 infants who received ASM before the EEG epoch analyzed).

**TABLE 6** ML models using aEEG<sup>a</sup> monitoring ( $n = 159$ )

| ML model                             | MCC   | AUC   | Sensitivity | Specificity | PPV  | NPV  |
|--------------------------------------|-------|-------|-------------|-------------|------|------|
| Quantitative-aEEG model              | 0.381 | 0.696 | 71.2        | 69.2        | 52.9 | 83.1 |
| Clinical and quantitative-aEEG model | 0.384 | 0.720 | 61.5        | 77.6        | 57.1 | 80.6 |

Abbreviations: ML, machine learning; MCC, Matthews correlation coefficient; AUC, area under the receiver-operating characteristic curve; PPV (NPV), positive (negative) predictive value.

<sup>a</sup>aEEG = two-channel EEG (F3-C3 and F4-C4) analysis included 159 infants.

close to the model reported previously by Sansevere et al.<sup>32</sup> (AUC 0.662) and outperformed the model reported by Jain et al.<sup>30</sup> In addition, we have shown that these models work equally well in the first hours after birth when important therapeutic decisions need to be made.

Given the increasing use of EEG monitoring in neonatal units worldwide, EEG background features have also been investigated to assess their ability to predict seizures and outcome and have been shown to outperform clinical parameters. Early abnormal EEG background (discontinuity and low voltage features) has been associated with seizure development.<sup>28,32,47,48</sup> Consistent with the literature, in our study, early low voltage to isoelectric EEG background was more frequently present in the seizure group compared to the non-seizure group (67.9% vs 20.2%,

$p$ -value <.001). Early SWC was absent in almost all infants with seizures (in our cohort only one infant with later seizures developed SWC before 12 h). In the non-seizure group, 16 infants had some evidence of early SWC, and all except one were diagnosed with mild encephalopathy. These findings might suggest that when early SWC is present, seizure occurrence is unlikely in the setting of mild HIE, and this could be reassuring for the bedside neonatologist. Quantitative-EEG analysis is an objective and reproducible analysis of the EEG background activity that does not require expert interpretation by a neurophysiologist. Low EEG power has been correlated previously with a poor neurodevelopmental outcome in neonates.<sup>49,50</sup> A more recent study has demonstrated that total EEG power < 10  $\mu\text{V}^2$  had a specificity of 98%, a sensitivity of

50%, and a PPV of 90% for seizure prediction.<sup>30</sup> We developed a seizure prediction model for infants with HIE, using a more extensive quantitative-EEG analysis including several power, discontinuity, spectral distribution, and inter-hemispheric connectivity features (sensitivity 69.8%, specificity 78.9%, PPV 61.7%, NPV 84.3%). A combination of all these features into a quantitative-EEG model had better prediction compared to the clinical model and a prediction similar to that of the qualitative-EEG model (the neurophysiologist's assessment). These findings suggest that quantitative-EEG analyses can be as accurate as expert interpretation to predict the likelihood of seizures and could augment the neurophysiology service, especially during out of hours periods, to guide the frequency of remote reviewing.

Sansevere et al. developed a combined neonatal seizure prediction model, including clinical information and qualitative-EEG background features, with an AUC (95% confidence interval [CI]) 0.830 (0.776–0.884).<sup>31</sup> In comparison, we developed different combined prediction models: clinical and qualitative-EEG model (AUC (95% CI) 0.721 (0.681 to 0.813)) and clinical and quantitative-EEG model (AUC (95% CI) 0.746 (0.700 to 0.833)). The addition of clinical parameters to qualitative-EEG and to quantitative-EEG analyses significantly improved the predictive ability of clinical model alone. These results emphasize the value of combining clinical and EEG analysis. However, a direct comparison with Sansevere's model is difficult because of several differences between the two studies: We used a more homogeneous population of term infants with HIE; timing of the EEG epochs included in our analysis was standardized to before 12h of age and before the emergence of electrographic seizures; qualitative analysis was based on the selected epoch and performed by an expert in neonatal EEG blinded to seizure occurrence; the proposed models were tested using a leave-one-baby-out cross-validation strategy, an unbiased estimator of the generalization performance. This testing process inherently incorporates variance of the models, a likely problem for small data sets such as this, which again minimizes potentially inflated estimates of performance when training and testing with the same data set.

Because many neonatal units worldwide use aEEG monitoring, we also developed a seizure-prediction model using two-channel EEG (aEEG – F3-C3 and F4-C4).<sup>51–53</sup> Although, the quantitative-aEEG model was less predictive than the multichannel EEG models, performance improved compared to the clinical model alone. Similarly, adding clinical information to the quantitative-aEEG analysis improved performance slightly.

The effect of ASM on EEG background was evident in this study by the difference in performance when excluding all infants treated with ASMs prior to the EEG,

resulting in models with a superior performance (clinical and quantitative-EEG model MCC 0.588). However, in clinical practice it would be difficult to exclude these infants, as clinical seizures are often treated before the EEG monitoring is started, especially for those infants that are outborn. Due to the relatively small sample size for this subgroup analysis ( $n = 131$  infants), more studies are warranted to account adequately for the effect of ASM.

Several limitations should be considered when interpreting these results. Although this study used a large study sample of infants with HIE, the numbers are small for ML analysis (<200 data points), limiting the ability to build complex ML models and potential bias toward overfitting. However, cross validation can limit this bias in performance assessment. In addition, there are different ways to combine ML models. The method chosen for this study was a late-stage fusion method, as we found this to be a better strategy than early-stage fusion for including all features (clinical + EEG) into one model. Due to missing data for blood gas analyses (ranging from 15%–37%), we used mean values to account for this, and included these parameters in the models, even if previous studies have demonstrated no significant change in the predictive value by adding them.<sup>27</sup> The presence of artifacts in the EEG trace is a limitation when performing quantitative analysis. However, for this analysis we used the first good quality hour of EEG recording, each hour was visually inspected, and artifacts were annotated and removed from the analysis. Unlike detecting seizure events, we are uncertain as to the specific EEG patterns that might be associated with the later development of seizures, other than low power, which is non-specific. Future work should focus on looking for specific patterns, waveforms, and features in the EEG background associated with later development of seizures. The confounding effect on the EEG background of therapeutic hypothermia and ASM administration in our analysis was reduced by adjusting the quantitative-EEG model analysis. The presented models were developed for term and near-term newborns with HIE (36–44 weeks of gestation); therefore new ML models would need to be developed to investigate the likelihood of seizures in other subgroups of newborn infants.

Despite these limitations, this is the first study to use a large cohort of infants with HIE from multiple European neonatal units with early, conventional multichannel EEG monitoring to predict the later development of seizures. The predictive performance of these models included a homogeneous cohort of infants with HIE, as this group represents the main cause of seizures in term newborns and the main reason for EEG monitoring in neonatal units.<sup>5,6</sup> The EEG background visual analysis and seizure detection were performed by experts in neonatal EEG interpretation, giving more strength to the qualitative-EEG analysis.

EEG analysis was based on an eight-channel EEG montage, which allowed us to assess cerebral activity across all cortical regions and to detect seizures that might be missed with more limited aEEG monitoring. In addition, a comprehensive quantitative-EEG analysis was included in the ML analysis as an objective and reproducible quantification of the cerebral activity. Predictive models were developed separately using the eight-channel EEG and two-channel EEG aEEG. These models are research focused and will therefore need real-time clinical validation.

## 5 | CONCLUSION

To summarize, early qualitative and quantitative-EEG features alone can predict infants who will later develop seizures in HIE, hours before seizure onset. Adding available clinical information augments the predictive value of these models. We demonstrated that ML analysis using quantitative EEG is as reliable as an expert (qualitative-EEG analysis) in predicting the likelihood of seizures. Automated EEG analysis may be useful to individualize the neurophysiology review frequency of continuous EEG monitoring, which may have a significant impact for units with less neurophysiology support. This study opens the door for the implementation of automated ML seizure-prediction algorithms, as objective bedside tools to identify which neonates with HIE are at highest risk of seizures.

### AUTHOR CONTRIBUTIONS

Andreea M. Pavel recruited participants and collected data, analyzed the data, drafted the initial manuscript, and revised the manuscript. John M O'Toole carried out the quantitative-EEG analysis and the ML modeling analysis. Jacopo Proietti carried out the qualitative-EEG analysis. Vicki Livingstone, Subhabrata Mitra, Eugene M. Dempsey, and Deirdre M. Murray were involved in conceptualization and design of the studies and coordinated and supervised data collection and analysis. Geraldine B. Boylan conceptualized, designed, and coordinated both studies; designed the EEG analysis protocol; and carried out the qualitative-EEG analysis. All authors reviewed and revised the manuscript.

### ACKNOWLEDGMENTS

The authors would like to thank to all the families for their participation and to all the clinical and research teams at each recruiting for all their hard work and dedication. Open access funding provided by IReL.

ANSER Consortium: INFANT Research Centre and Department of Pediatrics and Child Health, University College Cork, Cork, Ireland: Andreea M. Pavel, MD; Eugene

M. Dempsey, MD; Vicki Livingstone, PhD; Elena Pavlidis, MD; Liudmila Kharoshankaya, MD; Sean R. Mathieson, PhD; William P. Marnane, PhD; Gordon Lightbody, PhD; Jackie O'Leary, MSc; Mairead Murray, MSc; Jean Conway, MSc; Denis Dwyer, BSc; Andrey Temko, PhD; Taragh Kiely, MSc; Anthony C. Ryan, MD; Geraldine B. Boylan, PhD. Institute for Women's Health, University College London, London, UK: Janet M. Rennie, MD; Subhabrata Mitra, PhD. Utrecht Brain Center, University Medical Center Utrecht, Utrecht University, The Netherlands: Linda S. de Vries, PhD; Lauren C. Weeke, PhD; Mona C. Toet, PhD; Johanneke C. Harteman, PhD. Department of Neonatal Medicine and Division of Pediatrics, Department CLINTEC, Karolinska University Hospital, Karolinska Institutet, Stockholm, Sweden: Mikael Finder, PhD; Mats Blennow, PhD; Ingela Edqvist. Rotunda Hospital, Dublin, Ireland: Adrienne Foran, MD; Raga Mallika Pinnamaneni, MD; Jessica Colby-Milley, MSc. Royal London Hospital and Queen Mary University of London, London, UK: Divyen K. Shah, MD; Nicola Openshaw-Lawrence. Department of Clinical Neurophysiology, Great Ormond Street Hospital for Children NHS Trust, London, UK: Ronit M. Pressler, PhD. Homerton University Hospital NHS Foundation Trust, London, UK: Olga Kapellou, MD. Clinical Neurophysiology, University Medical Center Utrecht, Utrecht, The Netherlands: Alexander C. van Huffelen, PhD.

### FUNDING INFORMATION

This study was supported by: A Strategic Translational Award and an Innovator Award from the Wellcome Trust (098983 & 209325); Science Foundation Ireland (15/RI/3239 and 15/SIRG/3580); and the Health Research Board, Ireland (NEPTuNE CDA-2018-008). Funders had no role in the design and conduct of the study or in this analysis.

### CONFLICT OF INTEREST

All authors have no conflict of interest to disclose.

### ORCID

Andreea M. Pavel  <https://orcid.org/0000-0002-7233-3408>

John M. O'Toole  <https://orcid.org/0000-0001-9021-1183>

Mikael Finder  <https://orcid.org/0000-0002-3788-5493>

Geraldine B. Boylan  <https://orcid.org/0000-0003-0920-5291>

### REFERENCES

1. Lee AC, Kozuki N, Blencowe H, Vos T, Bahalim A, Darmstadt GL, et al. Intrapartum-related neonatal encephalopathy incidence and impairment at regional and global levels for 2010 with trends from 1990. *Pediatr Res.* 2013;74:50–72.
2. Gale C, Statnikov Y, Jawad S, Uthaya SN, Modi N. Neonatal brain injuries in England: population-based incidence derived

- from routinely recorded clinical data held in the National Neonatal Research Database. *Arch Dis Child Fetal Neonatal Ed.* 2018;103(4):F301–6.
3. Volpe JJ. *Neurology of the Newborn E-book.* Amsterdam: Elsevier Health Science, Publishing Company; 2008.
  4. Lynch NE, Stevenson NJ, Livingstone V, Murphy BP, Rennie JM, Boylan GB. The temporal evolution of electrographic seizure burden in neonatal hypoxic ischemic encephalopathy. *Epilepsia.* 2012;53(3):549–57.
  5. Tekgul H, Gauvreau K, Soul J, Murphy L, Robertson R, Stewart J, et al. The current etiologic profile and neurodevelopmental outcome of seizures in term newborn infants. *Pediatrics.* 2006;117(4):1270–80.
  6. Rennie JM, de Vries LS, Blennow M, Foran A, Shah DK, Livingstone V, et al. Characterisation of neonatal seizures and their treatment using continuous EEG monitoring: a multicentre experience. *Arch Dis Child Fetal Neonatal Ed.* 2018;104(5):F493–F451.
  7. Shankaran S, Laptook AR, Ehrenkranz RA, Tyson JE, McDonald SA, Donovan EF, et al. Whole-body hypothermia for neonates with hypoxic-ischemic encephalopathy. *N Engl J Med.* 2005;353(15):1574–84.
  8. Azzopardi DV, Strohm B, Edwards AD, Dyet L, Halliday HL, Juszczak E, et al. Moderate hypothermia to treat perinatal asphyxial encephalopathy. *N Engl J Med.* 2009;361(14):1349–58.
  9. Simbruner G, Mittal RA, Rohlmann F, Mueche R. Systemic hypothermia after neonatal encephalopathy: outcomes of the nEURO.Network RCT. *Pediatrics.* 2010;126(4):e771–8.
  10. Low E, Boylan GB, Mathieson SR, Murray DM, Korotchkova I, Stevenson NJ, et al. Cooling and seizure burden in term neonates: an observational study. *Arch Dis Child Fetal Neonatal Ed.* 2012;97(4):F267–72.
  11. Srinivasakumar P, Zempel J, Wallendorf M, Lawrence R, Inder T, Mathur A. Therapeutic hypothermia in neonatal hypoxic ischemic encephalopathy: electrographic seizures and magnetic resonance imaging evidence of injury. *J Pediatr.* 2013;163(2):465–70.
  12. Wusthoff CJ, Dlugos DJ, Gutierrez-Colina A, Wang A, Cook N, Donnelly M, et al. Electrographic seizures during therapeutic hypothermia for neonatal hypoxic-ischemic encephalopathy. *J Child Neurol.* 2011;26(6):724–8.
  13. Glass HC, Nash KB, Bonifacio SL, Barkovich AJ, Ferriero DM, Sullivan JE, et al. Seizures and magnetic resonance imaging-detected brain injury in newborns cooled for hypoxic-ischemic encephalopathy. *J Pediatr.* 2011;159(5):731–5.e1.
  14. Kharoshankaya L, Stevenson NJ, Livingstone V, Murray DM, Murphy BP, Ahearne CE, et al. Seizure burden and neurodevelopmental outcome in neonates with hypoxic-ischemic encephalopathy. *Dev Med Child Neurol.* 2016;58(12):1242–8.
  15. Nagarajan L, Palumbo L, Ghosh S. Neurodevelopmental outcomes in neonates with seizures: a numerical score of background encephalography to help prognosticate. *J Child Neurol.* 2010;25(8):961–8.
  16. Uria-Avellanal C, Marlow N, Rennie JM. Outcome following neonatal seizures. *Semin Fetal Neonatal Med.* 2013;18(4):224–32.
  17. Ronen GM, Buckley D, Penney S, Streiner DL. Long-term prognosis in children with neonatal seizures: a population-based study. *Neurology.* 2007;69(19):1816–22.
  18. Dzhala VI, Kuchibhotla KV, Glykys JC, Kahle KT, Swiercz WB, Feng G, et al. Progressive NKCC1-dependent neuronal chloride accumulation during neonatal seizures. *J Neurosci.* 2010;30(35):11745–61.
  19. Murray DM, Boylan GB, Ali I, Ryan CA, Murphy BP, Connolly S. Defining the gap between electrographic seizure burden, clinical expression and staff recognition of neonatal seizures. *Arch Dis Child Fetal Neonatal Ed.* 2008;93(3):F187–91.
  20. Scher MS, Alvin J, Gaus L, Minnigh B, Painter MJ. Uncoupling of EEG-clinical neonatal seizures after antiepileptic drug use. *Pediatr Neurol.* 2003;28(4):277–80.
  21. Malone A, Ryan CA, Fitzgerald A, Burgoyne L, Connolly S, Boylan GB. Interobserver agreement in neonatal seizure identification. *Epilepsia.* 2009;50(9):2097–101.
  22. Shellhaas RA, Chang T, Tsuchida T, Scher MS, Riviello JJ, Abend NS, et al. The American clinical neurophysiology Society's guideline on continuous electroencephalography monitoring in neonates. *J Clin Neurophysiol.* 2011;28(6):611–7.
  23. Pressler RM, Cilio MR, Mizrahi EM, Moshé SL, Nunes ML, Plouin P, et al. The ILAE classification of seizures and the epilepsies: modification for seizures in the neonate. Position paper by the ILAE task force on neonatal seizures. *Epilepsia.* 2021;62(3):615–28.
  24. Rennie J, Boylan G. Treatment of neonatal seizures. *Arch Dis Child Fetal Neonatal Ed.* 2007;92(2):F148–50.
  25. Perlman JM, Risser R. Can asphyxiated infants at risk for neonatal seizures be rapidly identified by current high-risk markers? *Pediatrics.* 1996;97(4):456–62.
  26. Andres RL, Saade G, Gilstrap LC, Wilkins I, Witlin A, Zlatnik F, et al. Association between umbilical blood gas parameters and neonatal morbidity and death in neonates with pathologic fetal acidemia. *Am J Obstet Gynecol.* 1999;181(4):867–71.
  27. Murray DM, Ryan CA, Boylan GB, Fitzgerald AP, Connolly S. Prediction of seizures in asphyxiated neonates: correlation with continuous video-electroencephalographic monitoring. *Pediatrics.* 2006;118(1):41–6.
  28. Glass HC, Wusthoff CJ, Shellhaas RA, Tsuchida TN, Bonifacio SL, Cordeiro M, et al. Risk factors for EEG seizures in neonates treated with hypothermia: a multicenter cohort study. *Neurology.* 2014;82(14):1239–44.
  29. Cornet MC, Pasupuleti A, Fang A, Gonzalez F, Shimotake T, Ferriero DM, et al. Predictive value of early EEG for seizures in neonates with hypoxic-ischemic encephalopathy undergoing therapeutic hypothermia. *Pediatr Res.* 2018;84(3):399–402.
  30. Jain SV, Mathur A, Srinivasakumar P, Wallendorf M, Culver JP, Zempel JM. Prediction of neonatal seizures in hypoxic-ischemic encephalopathy using electroencephalograph power analyses. *Pediatr Neurol.* 2017;67:64–70.
  31. Sanservere AJ, Kapur K, Peters JM, Fernandez IS, Lodenkemper T, Soul JS. Seizure prediction models in the neonatal intensive care unit. *J Clin Neurophys.* 2019;36(3):186–94.
  32. Macdonald-Laurs E, Sharpe C, Nespeca M, Rismanchi N, Gold JJ, Kuperman R, et al. Does the first hour of continuous electroencephalography predict neonatal seizures? *Arch Dis Child Fetal Neonatal Ed.* 2021;106(2):162–7.
  33. Struck AF, Tabaeizadeh M, Schmitt SE, Ruiz AR, Swisher CB, Subramaniam T, et al. Assessment of the validity of the 2HELPS2B score for inpatient seizure risk prediction. *JAMA Neurol.* 2020;77(4):500–7.

34. Schwalbe N, Wahl B. Artificial intelligence and the future of global health. *Lancet*. 2020;395(10236):1579–86.
35. Abbasi B, Goldenholz DM. Machine learning applications in epilepsy. *Epilepsia*. 2019;60(10):2037–47.
36. Pavel AM, Rennie JM, de Vries LS, Blennow M, Foran A, Shah DK, et al. A machine-learning algorithm for neonatal seizure recognition: a multicentre, randomised, controlled trial. *Lancet Child Adolesc Health*. 2020;4(10):740–9.
37. Jennekens W, Dankers F, Janssen F, Toet M, van der Aa N, Niemarkt H, et al. Effects of midazolam and lidocaine on spectral properties of the EEG in full-term neonates with stroke. *Eur J Paediatr Neurol*. 2012;16(6):642–52.
38. Obeid R, Tsuchida TN. Treatment effects on neonatal EEG. *J Clin Neurophysiol*. 2016;33(5):376–81.
39. Tsuchida TN, Wusthoff CJ, Shellhaas RA, Abend NS, Hahn CD, Sullivan JE, et al. American clinical neurophysiology society standardized EEG terminology and categorization for the description of continuous EEG monitoring in neonates: report of the American clinical neurophysiology society critical care monitoring committee. *J Clin Neurophysiol*. 2013;30(2):161–73.
40. Clancy RR, Legido A. The exact ictal and interictal duration of electroencephalographic neonatal seizures. *Epilepsia*. 1987;28(5):537–41.
41. O' Toole JM, Boylan GB. NEURAL: quantitative features for newborn EEG using Matlab. Preprint. 2017. <https://arxiv.org/abs/1704.05694>
42. O'Toole JM, Boylan GB, Lloyd RO, Goulding RM, Vanhatalo S, Stevenson NJ. Detecting bursts in the EEG of very and extremely premature infants using a multi-feature approach. *Med Eng Phys*. 2017;45:42–50.
43. Raurale SA, Nalband S, Boylan G, Lightbody G, O'Toole JM. Suitability of an inter-burst detection method for grading hypoxic-ischemic encephalopathy in newborn EEG. 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Berlin: IEEE; 2019. p. 4125–8.
44. Veronika Dorogush A, Ershov V, Gulin A. CatBoost: gradient boosting with categorical features support. arXiv. 2018; arXiv:1810.11363. <https://ui.adsabs.harvard.edu/abs/2018arXiv181011363V>. Accessed 01 Oct 2018.
45. Chicco D, Jurman G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*. 2020;21(1):6.
46. Satopää V, Baron J, Foster D, Mellers B, Tetlock P, Ungar L. Combining multiple probability predictions using a simple logit model. *Int J Forecast*. 2014;30:344–56.
47. Murray DM, Boylan GB, Fitzgerald AP, Ryan CA, Murphy BP, Connolly S. Persistent lactic acidosis in neonatal hypoxic-ischaemic encephalopathy correlates with EEG grade and electrographic seizure burden. *Arch Dis Child Fetal Neonatal Ed*. 2008;93(3):F183–6.
48. Laroia N, Guillet R, Burchfiel J, McBride MC. EEG background as predictor of electrographic seizures in high-risk neonates. *Epilepsia*. 1998;39(5):545–51.
49. Bell AH, McClure BG, Hicks EM. Power spectral analysis of the EEG of term infants following birth asphyxia. *Dev Med Child Neurol*. 1990;32(11):990–8.
50. Korotchikova I, Stevenson NJ, Walsh BH, Murray DM, Boylan GB. Quantitative EEG analysis in neonatal hypoxic ischaemic encephalopathy. *Clin Neurophysiol*. 2011;122(8):1671–8.
51. Shah NA, Van Meurs KP, Davis AS. Amplitude-integrated electroencephalography: a survey of practices in the United States. *Am J Perinatol*. 2015;32(8):755–60.
52. Boylan G, Burgoyne L, Moore C, O'Flaherty B, Rennie J. An international survey of EEG use in the neonatal intensive care unit. *Acta Paediatr*. 2010;99(8):1150–5.
53. Mescall S, Dhamodaran M, Clarke P, Ponnusamy V. Current availability of cerebral function monitoring and therapeutic hypothermia equipment in UK neonatal units and networks. *Arch Dis Childhood Fetal Neonatal Ed*. 2021;106(2):225–7.

## SUPPORTING INFORMATION

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

**How to cite this article:** Pavel AM, O'Toole JM, Proietti J, Livingstone V, Mitra S, Marnane WP, et al. Machine learning for the early prediction of infants with electrographic seizures in neonatal hypoxic-ischemic encephalopathy. *Epilepsia*. 2023;64:456–468. <https://doi.org/10.1111/epi.17468>