Leveraging electronic medical record-embedded standardised electroencephalogram reporting to develop neonatal seizure prediction models: a retrospective cohort study

Jillian L McKee, MD,

Michael C Kaufman, MS,

Alexander K Gonzalez, MS,

Mark P Fitzgerald, MD,

Shavonne L Massey, MD,

France Fung, MD,

Sudha K Kessler, MD,

Stephanie Witzman, BA,

Nicholas S Abend, MD,

Ingo Helbig, MD

(J L McKee MD, M C Kaufman MS, M P Fitzgerald MD, S L Massey MD, F Fung MD, S K Kessler MD, S Witzman BA, Prof N S Abend MD, I Helbig MD), **The Epilepsy NeuroGenetics Initiative** (J L McKee, M C Kaufman, M P Fitzgerald, S L Massey, I Helbig), **and Department of Biomedical and Health Informatics** (M C Kaufman, A K Gonzalez MS, I Helbig), **Children's Hospital of Philadelphia, Philadelphia, PA, USA; Department of Neurology** (J L McKee, M P Fitzgerald, S L Massey, F Fung, S K Kessler, Prof N S Abend, I Helbig) **and Department of Anesthesia and Critical Care Medicine** (Prof N S Abend), **Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA**

Summary

Background—Accurate prediction of seizures can help to direct resource-intense continuous electroencephalogram (CEEG) monitoring to neonates at high risk of seizures. We aimed to use data from standardised EEG reports to generate seizure prediction models for vulnerable neonates.

Methods—In this retrospective cohort study, we included neonates who underwent CEEG during the first 30 days of life at the Children's Hospital of Philadelphia (Philadelphia, PA, USA). The hypoxic ischaemic encephalopathy subgroup included only patients with CEEG data during the first 5 days of life, International Classification of Diseases, revision 10, codes for hypoxic ischaemic encephalopathy, and documented therapeutic hypothermia. In January, 2018, we implemented a novel CEEG reporting system within the electronic medical record (EMR) using common data elements that incorporated standardised terminology. All neonatal CEEG data from Jan 10, 2018, to Feb 15, 2022, were extracted from the EMR using age at the time of CEEG. We developed logistic regression, decision tree, and random forest models of neonatal seizure prediction using EEG features on day 1 to predict seizures on future days.

Findings—We evaluated 1117 neonates, including 150 neonates with hypoxic ischaemic encephalopathy, with CEEG data reported using standardised templates between Jan 10, 2018, and Feb 15, 2022. Implementation of a consistent EEG reporting system that documents discrete and standardised EEG variables resulted in more than 95% reporting of key EEG features. Several EEG features were highly correlated, and patients could be clustered on the basis of specific features. However, no simple combination of features adequately predicted seizure risk. We therefore applied computational models to complement clinical identification of neonates at high risk of seizures. Random forest models incorporating background features performed with classification accuracies of up to 90% (95% CI 83–94) for all neonates and 97% (88–99) for neonates with hypoxic ischaemic encephalopathy; recall (sensitivity) of up to 97% (91–100) for all neonates and 100% (100–100) for neonates with hypoxic ischaemic encephalopathy; and precision (positive predictive value) of up to 92% (84–96) in the overall cohort and 97% (80–99) in neonates with hypoxic ischaemic encephalopathy.

Interpretation—Using data extracted from the standardised EEG report on the first day of CEEG, we predict the presence or absence of neonatal seizures on subsequent days with classification performances of more than 90%. This information, incorporated into routine care, could guide decisions about the necessity of continuing EEG monitoring beyond the first day, thereby improving the allocation of limited CEEG resources. Additionally, this analysis shows the benefits of standardised clinical data collection, which can drive learning health system approaches to personalised CEEG use.

Introduction

Neonatal seizures are common and contribute to morbidity and mortality.[1] Neonates with hypoxic ischaemic encephalopathy have a high incidence of seizures (about 30%),[2–4] which are identified mostly by electroencephalogram (EEG) only and would not be identified by clinical observation alone. Additionally, among neonates with hypoxic ischaemic encephalopathy, seizures have been associated with an increased risk of subsequent neurobehavioural problems and epilepsy.[2,4–7] Because of the high risk of seizures in this population, guidelines recommend that neonates with hypoxic ischaemic encephalopathy undergo continuous EEG (CEEG) monitoring during the periods of therapeutic hypothermia and return to normothermia, which often last 4–5 days.[8,9] However, this practice is resource intensive and not feasible for all neonates who might benefit[10] because many neonates receive care in neonatal intensive care units without CEEG capability and approaches to

remote CEEG are not widely available.[11] Furthermore, CEEG is not entirely benign, as long-term electrode placement can cause skin breakdown[12] and the necessary wires and head-wrap can interfere with parent–infant physical interaction and feeding.[13] Therefore, it would be helpful to predict each individual's risk of seizure to optimise resource allocation and minimise unnecessary medical procedures by targeting CEEG to neonates at the highest risk of seizures.

Previous studies have shown that the prediction of neonatal seizures is complex, and clinical and EEG data generally do not predict seizures well. Clinical features alone are not predictive of seizures in neonates.[3,14] EEG studies have found that although normal backgrounds correctly predict the absence of seizures,[15] an abnormal background does not accurately predict the presence of seizures.[3,15–17] Seizure prediction models based on EEG features determined by manual review of EEG segments,[18–21] review of EEG reports,[16,17] or direct computational analysis of EEG tracings[22] have been limited by small sample sizes or the short windows of review,[23] and only a few have been tested in neonates.[16,17,19–21] Due to the need for manual chart or EEG review, the existing models cannot easily be incorporated into routine clinical care.

We implemented a standardised reporting template for all clinical EEG reports derived from terminology published by the American Clinical Neurophysiology Society (ACNS),[24] and have previously shown that this reporting system is acceptable to electroencephalographers.[25] In this study, we aimed to determine whether this standardised template leads to complete reporting of the recommended EEG features, and we aimed to develop neonatal seizure prediction models based on data extracted from these reports to optimise CEEG use among neonates, including the subset with hypoxic ischaemic encephalopathy.

Methods

Study design and participants

This retrospective cohort study was done at the Children's Hospital of Philadelphia (Philadelphia, PA, USA). All neonatal CEEG data from Jan 10, 2018, to Feb 15, 2022, were extracted from the electronic medical record (EMR) using age at the time of CEEG and the EEG study common data elements. Neonates were included in this cohort if they underwent CEEG during the first 30 days of life. The hypoxic ischaemic encephalopathy subgroup included only patients with CEEG data during the first 5 days of life, International Classification of Diseases, revision 10, codes for hypoxic ischaemic encephalopathy, and documented therapeutic hypothermia. The patient list was cross-validated against an independently curated list kept by the Critical Care EEG Service at the study site, with manual chart review of selected patients to validate the EMR algorithm (appendix p 1). Other indications for neonatal CEEG included cardiac surgery for con genital heart disease, congenital diaphragmatic hernia repair, and clinical concern for seizure.

This study met exemption criteria per the Institutional Review Board at the Children's Hospital of Philadelphia (Philadelphia, PA, USA), and, as such, patient consent was not required.

Procedures

Patients were managed by neonatology and neurocritical care consultation services. CEEG was performed following an institutional CEEG pathway consistent with ACNS recommendations[8] and generally lasted 2 days for most neonates, 4–5 days for neonates with hypoxic ischaemic encephalopathy, and at least 24 h after the last seizure for neonates with seizures. Prophylactic anti-seizure medications were not administered, but some neonates with clinically evident seizures before CEEG initiation were receiving antiseizure medications. All neonates with electrographic seizures were treated with appropriate antiseizure medications.

CEEGs were interpreted by paediatric electroencephalographers who were knowledgeable about the ACNS standardised terminology.[24] In January, 2018, the Children's Hospital of Philadelphia implemented a novel EEG reporting system based on common data elements in the EMR that incorporated this terminology. The neonatal template included embedded descriptions of continuity, variability, reactivity, voltage, graphoelements, epileptiform transients, seizures, and an overall impression (appendix p 1). According to ACNS guidelines, electrographic seizures were defined as repetitive, evolving patterns with a minimum voltage of 2 mV peak-to-peak and a duration of at least 10 s.[24] Details regarding the reporting system have been published previously.[25] The electroencephalographers completed a standardised form that both stored the data as common data elements in the EMR and generated a clinical report. Thus, the information stored in the common data elements is the exact data used for clinical decision making.

Data processing

The diagnostic codes, demographic data, and EEG template common data elements extracted from the EMR were exported using Clarity, an SQL database, and analysed in R (version 4.1.2) using RStudio (version 2021.09.2+382). From these raw data, key predictive features from the first day of CEEG monitoring were selected and mapped onto binary outcomes representing "normal" and "abnormal" for ease of modelling. For example, the continuity feature included the levels "low voltage suppressed", "excessive discontinuity", and "burst suppression", which were mapped to "abnormal", and "normal discontinuity" and "normal continuity", which were in turn mapped to "normal". The outcome variable was defined as seizures on subsequent CEEG days, occurring on day 2 or later of the CEEG session.

Correlation and clustering analyses

Cross-correlation plots were created for all EEG features using native R functions and the corrplot package. Pearson cross-correlation values and 95% CIs were calculated. Features and patients were clustered using the ComplexHeatmap package,[26] separating patients with and without any seizures and using complete linkage clustering and a Euclidean distance measure.

Model building

Various frameworks can be used for prediction models, and we sought to identify the benefits and weaknesses of various methods to optimise the prediction of future seizures.

Our primary aim was to build a seizure prediction model to determine which neonates, and particularly which neonates with hypoxic ischaemic encephalopathy, would ultimately have seizures, based on EEG features during the first 24 h of CEEG. We compared logistic regression models, decision trees, and random forest algorithms, representing a shift from standard epidemiological methods to machine learning methods. Decision trees, which are supervised machine learning algorithms useful for solving classification problems with categorical variables, are simple, intuitive, fast, and relatively robust to outliers and missing data. An example decision tree is shown in the appendix (p 4). One preferred feature of these models is the ability to identify patients at low and high risk of seizures by examining the leaves of the tree. Random forest models are built using ensembles of many decision trees. The assumption underlying these models is that large numbers of uncorrelated trees operating as a group will outperform any constituent model. This prevents overfitting, which can occur with decision trees, and improves the stability of the final model.

Given the relative rarity of our outcome of interest, we tuned many of our models to be more sensitive to individuals with seizures by increasing the weight of samples with subsequent seizures. See the appendix (pp 5–6) for details of the weighting process. Logistic regression models were built using the caret R package. Random forest models were built using the randomForest and ranger packages as well as the H2O platform (H2O.ai, R Interface for H2O, R package, version 3.10.0.8). When evaluating the performance of our classifiers, we used accuracy, precision, recall, F1, area under the curve (AUC), area under the precision–recall curve (AUCPR), and Cohen's κ scores. Given our imbalanced dataset, we sought to maximise F1 and AUCPR over AUC and accuracy.[27] Variable importance, which displays the relative influence of each feature, was determined within our H2O models through feature inclusion in a split as well as the squared error reduction at each split in which that feature was included. All models were developed with separate training and testing data as well as cross-validation within the training data. We used bootstrapping to determine 95% CIs for model performance metrics. If neonates had two separate EEG recording periods separated by at least 24 h off EEG, then these were treated as two separate sessions for modelling.

Patients with missing data for any of the key variables were not included in the models. Missing data were excluded for three reasons. First, to facilitate comparisons across models, we only included individuals with complete data because not all models used can accommodate missing data. Second, working with complete data allowed us to better characterise the effect of each feature on model performance. Finally, based on data review and clinical experience with the templated EEG reports, the omitted clinical features are often normal and thus the exclusion of these individuals is not likely to falsely increase model accuracy. However, including individuals with incomplete data could be informative, potentially further improving the performance of models that can accommodate missing data.

Statistical analysis

We used Fisher's exact test to determine the significance of differences in proportions of categorical EEG features between groups. We used the Wilcoxon rank-sum test, a non-

parametric test, to compare numerical values between the overall neonatal group and the hypoxic ischaemic encephalopathy subgroup. In all cases, results were considered to be significant if p values were less than 0·05.

All code and the de-identified dataset are available online. All analyses were done with R (version 4.1.2).

Role of the funding source

The funders of the study had no role in study design, the collection, analysis, and interpretation of data, the writing of the report, or in the decision to submit the paper for publication.

Results

Between Jan 10, 2018, and Feb 15, 2022, the common data element-based novel documentation system was used to report more than 42 000 EEGs (figure 1A; appendix p 2), including 1117 neonates in the intensive care unit. For neonates, ACNS-defined key neonatal EEG variables (presence or absence of EEG seizures, continuity, variability, reactivity, voltage, graphoelements, epileptiform transients, and overall EEG impression) were recorded in more than 95% of reports on the first day of CEEG (figure 1C; appendix p 7). 114 neonates had two separate EEG recording periods separated by at least 24 h off EEG, and these were treated as two separate sessions for modelling. 277 neonates, 30 of whom had hypoxic ischaemic encephalopathy, were missing data for at least one key variable and were therefore excluded from models.

In addition to an overall high completion rate, the completeness of EEG reports improved over time (figure 1B). Neonatal EEGs reported in 2018, during the first 6 months of template use, had an average completion rate of 80%, whereas EEGs reported in 2021–22, during the most recent 6 months of template use, had a significantly higher average completion rate of 98% (p<0·0001). Therefore, implementation of EEG templates provided complete data that could be used to build prediction models to inform clinical care.

The cohort of neonates who underwent CEEG during the first 30 days of life included 1117 individuals, 150 (13·4%) of whom had hypoxic ischaemic encephalopathy managed with therapeutic hypothermia and constituted the hypoxic ischaemic encephalopathy subgroup for the seizure prediction analyses. The hypoxic ischaemic encephalopathy cohort is of clinical interest and serves as an example of a particularly vulnerable population that could benefit from improved intervention. Demographic features are presented in the table. The median age at CEEG initiation was 4 days (IQR 2–7) for the overall cohort and was significantly younger in the hypoxic ischaemic encephalopathy subgroup at 1 day (1–1, p<0·0001; table; appendix p 2). The median duration of CEEG was 3 days (IQR 2–4) for the overall cohort and was significantly longer at 5 days (4–6, p<0·0001; table; appendix p 2) in the hypoxic ischaemic encephalopathy subgroup.

On the first day of recording, the frequency of EEG feature abnormalities was higher in neonates in the hypoxic ischaemic encephalopathy subgroup than in the overall cohort

(p<0·0001 for voltage, p=0·0003 for variability, p=0·032 for reactivity, and p=0·0001 for graphoelements; figure 1C; appendix p 7). Over the course of CEEG, neonates in the hypoxic ischaemic encephalopathy subgroup were also more likely to have at least one seizure across all days of recording than the overall cohort, although this difference did not reach significance (49 [32·7%] of 150 *vs* 342 [26·0%] of 1313; odds ratio 1·4 [95% CI 0·98–2·1], p=0·060). The time to first seizure from CEEG initiation was similar between the overall cohort and hypoxic ischaemic encephalopathy subgroup (figure 1D).

Several EEG features were correlated for the overall cohort and the hypoxic ischaemic encephalopathy subgroup, including overall EEG impression and continuity (0·75 [95% CI 0·72–0·77] *vs* 0·78 [0·71–0·84]), variability and reactivity (0·48 [0·43–0·52] *vs* 0·62 [0·51–0·72]), voltage and variability (0·40 [0·35–0·44] *vs* 0·62 [0·51–0·71]), and voltage and reactivity (0·39 [0·35–0·44] *vs* 0·54 [0·41–0·65]; figure 2A, C). Furthermore, although there were overall differences in abnormal feature representation between the seizure and no seizure groups, there was a large degree of heterogeneity, and the presence or absence of seizures could not be predicted by feature clustering alone (figure 2B, D). In summary, in addition to being sparsely represented across the cohort, many EEG features were correlated, and easily identifiable patterns in the data were not predictive of future seizures.

Models built with logistic regression predicted subsequent seizures with accuracy of up to 84% (95% CI 78–89) and AUCPR of 0·54 (95% CI 0·35–0·71) in the overall cohort and accuracy of 77% (50–91) and AUCPR of 0·57 (0·34–0·86) in the hypoxic ischaemic encephalopathy subgroup (figures 3, 4).

A basic decision tree model did not outperform logistic regression, with accuracies of 82% (95% CI 75–86) for the overall cohort and 73% (45–86) for those with hypoxic ischaemic encephalopathy, and AUCPR of 0·36 (95% CI 0·23–0·60) for the overall cohort and 0·46 (0·28–0·73) for the hypoxic ischaemic encephalopathy subgroup (figures 3, 4). Thus, although decision trees are intuitive and could identify groups of patients at high and low risk of subsequent seizures more readily than logistic regression models, they had suboptimal performance.

Untuned random forest models performed our classification task with accuracies similar to the regression models. However, tuned models that more heavily weighted the subsequent seizure class helped the model to identify these patients with much greater success, resulting in recall of up to 97% (95% CI 91–100) in the overall cohort and 100% (100–100) in the hypoxic ischaemic encephalopathy subgroup. AUCPR values were up to 0·97 (95% CI 0·94–0·98) for the overall cohort and 0·98 (0·88–1·0) for the hypoxic ischaemic encephalopathy subgroup. This class weighting came at a cost of lower prediction accuracy for the null class (those who do not have seizures) and increased the false positive rate for the most heavily weighted models up to 53% (95% CI 26–90) for the overall cohort and 38% (13–72) for the hypoxic ischaemic encephalopathy subgroup. However, the accuracies of these models were still up to 90% (95% CI 83–94) for the overall cohort and 97% (88–99) for the hypoxic ischaemic encephalopathy subgroup, with precision (positive predictive value) of up to 92% (84–96) for the overall cohort and 97% (80–99) for the subgroup (figures 3, 4).

Discussion

Using neonatal EEG reports collected over a 4-year period at a large tertiary paediatric centre, we found that neonatal seizures on subsequent days after initial recording are highly predictable. Our results provide the basis for rational CEEG use, particularly in limited settings. For example, our findings suggest that a centre without the ability to monitor neonates with hypoxic ischaemic encephalopathy with CEEG through re-warming (4–5 days) could perform 24 h of CEEG and be assured that subsequent seizures would not be missed in neonates identified as being at low risk.

Our study had four main findings. First, we showed that use of a standardised EEG reporting template in the EMR improved adherence with use of recommended neonatal EEG terminology,[24] increasing over time from 80% to 98%. This result speaks to a dynamic interaction between tool implementation and provider uptake: even though compliant use of the new format was not monitored or reinforced, EEG readers slowly gravitated towards more complete reporting. This example shows how implementation of novel technologies can provide a subtle transformative effect in areas of health care that require exact and standardised reporting. Standardised EEG reporting might enable better understanding of the EEG data by clinical teams and better assessment for EEG changes over time, while also facilitating acquisition of EEG data required for research and quality improve ment initiatives.

Second, we showed that clinical and EEG data available from the first day of CEEG are highly predictive of seizures on the subsequent days. Using a random forest model, AUCPR values were 0·97 (95% CI 0·94–0·98) for the overall cohort and 0·98 (0·88–1·00) for the hypoxic ischaemic encephalopathy subgroup. Thus, for neonates with heterogeneous aetiologies as well as a smaller more homogeneous cohort, the considerable data obtained as part of routine care could be used to make reliable predictions about the occurrence of future events. CEEG is resource intense, and longer durations of CEEG, especially for patients at low risk of seizures, have substantial incremental cost-effectiveness ratios.[10] Thus, approaches to provide broad 1-day screening CEEG, but longer CEEG in patients at high risk of seizures, might yield a higher value strategy enabling more widespread implementation.

Third, we found that machine learning methods surpass the accuracy of conventional statistical approaches for seizure prediction in neonates using data from standardised neonatal EEG reports. This finding and the lack of predictive power of single EEG features can be attributed to the complexity inherent in EEG reports. Many EEG features are highly correlated, which represents an ongoing challenge for traditional statistical approaches. Tools such as random forest models are less susceptible to such confounding structures and can therefore generate more accurate predictions. Additionally, we used class weighting to optimise model performance, which maximised recall but also increased the false positive rates. For clinical applications, this represents a strength of this model because it can detect the neonates who will have seizures with very high sensitivity, while still identifying a sufficiently large proportion of neonates at low risk of seizures.

Fourth, use of standardised EEG reporting templates for 4 years as part of clinical reporting generated sufficient data to develop a model that could predict neonatal seizures on subsequent days with high accuracy. This finding speaks to the overall power of learning health system approaches in child neurology.[28] Basic standardisation and tracking could advance knowledge in a manner that can drive changes to clinical care, which can be further assessed and refined over time. The strong predictive power of these data supports the value of clinically derived information, which might be considered only indirect at first. Each EEG report represents the expert assessment of an electroencephalo grapher, thereby allowing us to use a data resource that is highly informative and valuable while also saving time and resources. Although multi-rater scoring with formal processes to derive consensus might represent the gold standard for EEG interpretation in research studies, such approaches are resource intense and not feasible for large-scale real-world investigations or learning health system approaches involving CEEG data. Our data indicate the potential for standardised EMR-based approaches to advance knowledge in an efficient manner.

Inferences about underlying causes and biology are limited in machine learning approaches in contrast to single-feature predictions or logistic regression. Although simple decision trees can be visualised and followed to guide clinical practice, random forest models do not have a simple or intuitive visual representation. Thus, we created an online calculator that generates predictions based on our model. We intend to update the model quarterly with the most recent data, following the paradigm of a learning health system.[28] We recognise the model cannot be externally validated at this time due to the scarcity of large standardised CEEG databases, and the accuracy at centres with differing patient populations is unknown. Furthermore, as we recognise that black-box models can sometimes be difficult to adopt into clinical practice, we explored the feature importance and found EEG features important within the model corresponded with our clinical experience.

In the future, the efficacy of seizure prediction could be improved further by considering additional data already existing in the EMR beyond the demographic features included in our current models. For example, clinical variables in combination with EEG features have previously resulted in improved model performance. For a seizure prediction model built using Cox's proportional hazards regression, the AUC improved from 76% to 83% when clinical features, including gestational age, EEG indication, and aetiology or therapies, were added to EEG-based prediction alone.[16] Features such as gestational age, Apgar scores, and diagnosis, and phenotypic features including human phenotype ontology codes,[29] physical examination signs, medication administration, and laboratory data are increasingly standardised within the EMR. As these data accrue over time, we anticipate that it will be feasible and beneficial to incorporate additional variables into our predictive models.

In summary, we have built the first high-performing prediction model for neonatal seizures occurring after the first day of CEEG using automated EEG docu mentation, achieving accuracies of more than 90%. The use of data from routine clinical care alone, without prior assumptions, was sufficient for meaningful seizure predictions in neonates at high risk of seizures. If validated, this model could enable more targeted use of limited CEEG resources by reducing CEEG duration among patients at low risk of seizures after the initial

day of CEEG. Incorporation of accurate seizure prediction into real-time clinical care could improve the quality and efficiency of care for neonates with critical illness.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Data sharing

Data in a de-identified format will be made available by request to the corresponding author. A cleaned version of the dataset and all code is available at https://github.com/helbig-lab/neonate-seizure-prediction.

References

1. Lanska MJ, Lanska DJ, Baumann RJ, Kryscio RJ. A population-based study of neonatal seizures in Fayette County, Kentucky. Neurology 1995; 45: 724–32. [PubMed: 7723962]

2. Glass HC, Nash KB, Bonifacio SL, et al. Seizures and magnetic resonance imaging-detected brain injury in newborns cooled for hypoxic-ischemic encephalopathy. J Pediatr 2011; 159: 731–35. [PubMed: 21839470]

3. Glass HC, Wusthoff CJ, Shellhaas RA, et al. Risk factors for EEG seizures in neonates treated with hypothermia: a multicenter cohort study. Neurology 2014; 82: 1239–44. [PubMed: 24610326]

4. Wusthoff CJ, Dlugos DJ, Gutierrez-Colina A, et al. Electrographic seizures during therapeutic hypothermia for neonatal hypoxicischemic encephalopathy. J Child Neurol 2011; 26: 724–28. [PubMed: 21447810]

5. Shellhaas RA, Wusthoff CJ, Numis AL, et al. Early-life epilepsy after acute symptomatic neonatal seizures: a prospective multicenter study. Epilepsia 2021; 62: 1871–82. [PubMed: 34212365]

6. Glass HC, Numis AL, Gano D, Bali V, Rogers EE. Outcomes after acute symptomatic seizures in children admitted to a neonatal neurocritical care service. Pediatr Neurol 2018; 84: 39–45. [PubMed: 29886041]

7. 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: 465–70. [PubMed: 23452588]

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: 611–17. [PubMed: 22146359]

9. Pressler RM, Cilio MR, Mizrahi EM, 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: 615–28. [PubMed: 33522601]

10. Abend NS, Topjian AA, Williams S. How much does it cost to identify a critically ill child experiencing electrographic seizures? J Clin Neurophysiol 2015; 32: 257–64. [PubMed: 25626776]

11. Fitzgerald MP, Massey SL, Fung FW, et al. Expanding access to continuous EEG monitoring in neonatal intensive care units. J Clin Neurophysiol 2021; 38: 525–29. [PubMed: 32541608]

12. Pasupuleti A, Amling J, Chang T, Scafidi J, Tsuchida T. Skin integrity during prolonged EEG recording in hospitalized neonatal and pediatric patients (P3.247). Neurology 2016;86 (suppl 16): 247.

13. Bonner O, Beardsall K, Crilly N, Lasenby J. 'There were more wires than him': the potential for wireless patient monitoring in neonatal intensive care. BMJ Innov 2017; 3: 12–18.

14. 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: 41–46. [PubMed: 16818547]

15. Glauser TA, Clancy RR. Adequacy of routine EEG examinations in neonates with clinically suspected seizures. J Child Neurol 1992;7: 215–20. [PubMed: 1573242]

16. Sansevere AJ, Kapur K, Peters JM, Fernández IS, Loddenkemper T, Soul JS. Seizure prediction models in the neonatal intensive care unit. J Clin Neurophysiol 2019; 36: 186–94. [PubMed: 30882530]

17. Worden LT, Chinappen DM, Stoyell SM, et al. The probability of seizures during continuous EEG monitoring in high-risk neonates. Epilepsia 2019; 60: 2508–18. [PubMed: 31745988]

18. Fung FW, Jacobwitz M, Parikh DS, et al. Development of a model to predict electroencephalographic seizures in critically ill children. Epilepsia 2020; 61: 498–508. [PubMed: 32077099]

19. Laroia N, Guillet R, Burchfiel J, McBride MC. EEG background as predictor of electrographic seizures in high-risk neonates. Epilepsia 1998; 39: 545–51. [PubMed: 9596208]

20. 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. [PubMed: 28062149]

21. Benedetti GM, Vartanian RJ, McCaffery H, Shellhaas RA. Early electroencephalogram background could guide tailored duration of monitoring for neonatal encephalopathy treated with therapeutic hypothermia. J Pediatr 2020; 221: 81–87. [PubMed: 32222256]

22. Li F, Liang Y, Zhang L, et al. Transition of brain networks from an interictal to a preictal state preceding a seizure revealed by scalp EEG network analysis. Cogn Neurodynamics 2019; 13: 175–81.

23. Schulze-Bonhage A, Feldwisch-Drentrup H, Ihle M. The role of high-quality EEG databases in the improvement and assessment of seizure prediction methods. Epilepsy Behav 2011;22 (suppl 1): S88–93. [PubMed: 22078525]

24. Tsuchida TN, Wusthoff CJ, Shellhaas RA, 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: 161–73. [PubMed: 23545767]

25. Witzman S, Massey SL, Kessler S, et al. Acceptability of standardized EEG reporting in an electronic health record. J Clin Neurophysiol 2020; 37: 455–61. [PubMed: 31688354]

26. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 2016; 32: 2847–49. [PubMed: 27207943]

27. Saito T, Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLoS One 2015; 10: e0118432. [PubMed: 25738806]

28. Grinspan ZM, Patel AD, Shellhaas RA, et al. Design and implementation of electronic health record common data elements for pediatric epilepsy: foundations for a learning health care system. Epilepsia 2021; 62: 198–216. [PubMed: 33368200]

29. Lewis-Smith D, Galer PD, Balagura G, et al. Modeling seizures in the Human Phenotype Ontology according to contemporary ILAE concepts makes big phenotypic data tractable. Epilepsia 2021;62: 1293–305. [PubMed: 33949685]

Research in context

Evidence before this study

We searched the literature on electroencephalogram (EEG)-based seizure prediction in neonates in PubMed from Jan 1, 1946, to June 1, 2022, using combinations of the keywords "seizure", "prediction", "EEG", "neonatal", and "hypoxic-ischemic encephalopathy". We used no language restrictions. Previous studies relied on manual review of EEG reports to forecast seizures in neonates using regression models. These studies were limited in sample size because they required manual review of reports and manual data entry.

No studies were identified using automated collection of EEG data from routine care, and none used machine learning-based modelling techniques.

Added value of this study

We built seizure prediction models based on standardised EEG features reported in the electronic medical record (EMR) that could predict seizures in neonates, and particularly neonates with hypoxic ischaemic encephalopathy, with greater than 90% accuracy. Furthermore, these models could be tuned to not miss seizures, performing with recall (sensitivity) of up to 97% (95% CI 91–100) in the overall neonatal cohort and 100% (100–100) in neonates with hypoxic ischaemic encephalopathy, while still maintaining precision (positive predictive value) of up to 92% (95% CI 84–96) in the overall cohort and 97% (80–99) in neonates with hypoxic ischaemic encephalopathy.

Previous studies have built seizure prediction models using EEG data, but most have used features derived from manual scoring of EEG tracings or computational analysis of the raw EEG recordings. Although these studies are informative, they are not easily scalable for incorporation into routine clinical practice. To our knowledge, this is the first study reporting a seizure prediction model based on standardised reports already documented in the EMR that can be used for clinical decision support to improve care for critically ill neonates.

Implications of all the available evidence

Continuous EEG (CEEG) monitoring is currently the standard of care for critically ill children at risk of seizures. Although effective for seizure detection, CEEG is resource intensive, especially when used for long durations in patients at low risk of seizures, and can have physical and psychosocial consequences, such as skin breakdown and impediments to parent–infant physical interaction. Accurately predicting which neonates are likely to have seizures after an initial shorter period of monitoring would help to allocate CEEG resources towards neonates with the highest risk of seizures and avoid unnecessary use of limited CEEG resources in neonates at low risk of seizures. Furthermore, the ability to directly extract these predictors from standardised clinical reporting in the EMR will allow for automated predictions and dashboard development for use at scale in real-time clinical care.
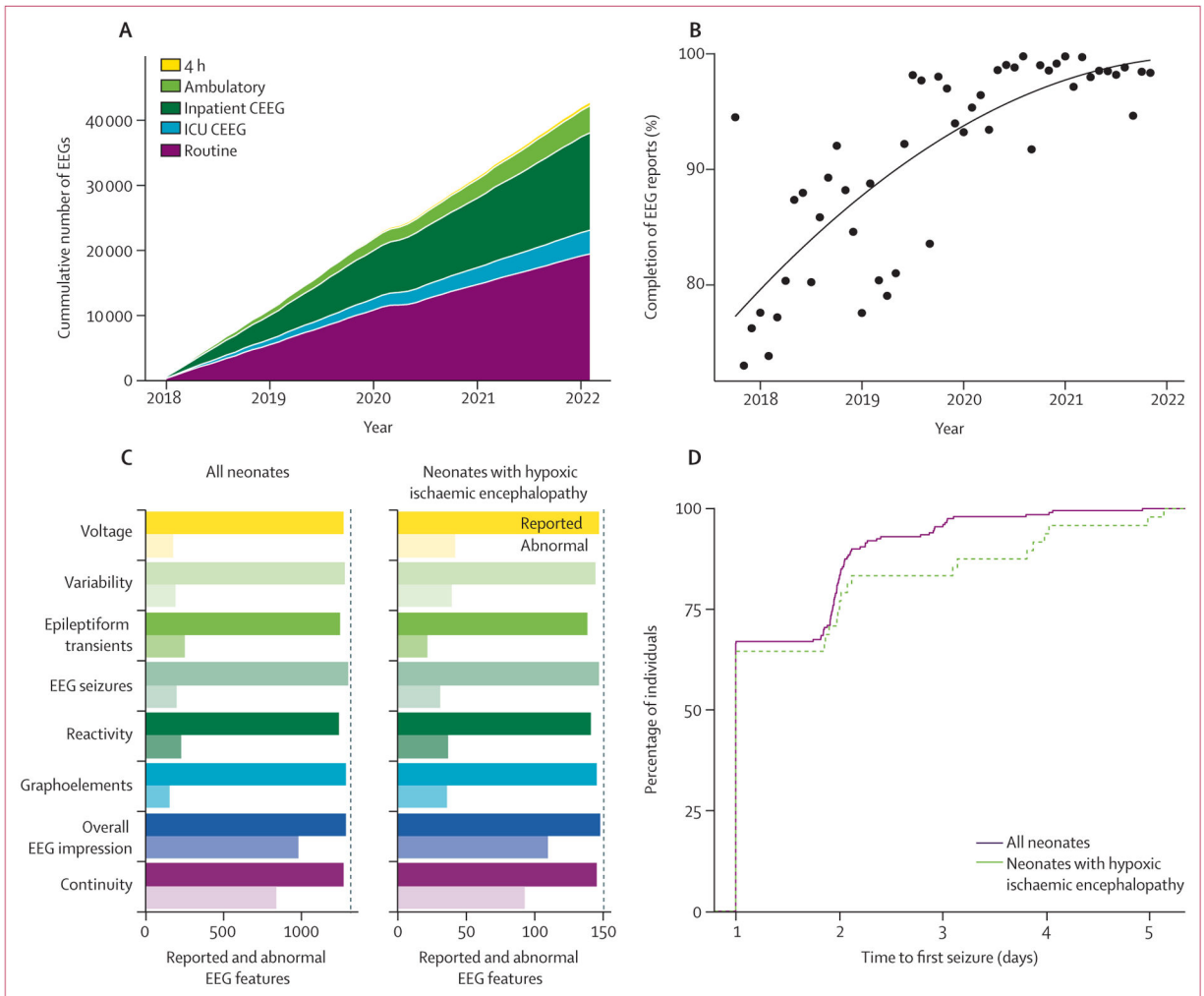
Figure 1: EEG data accrual, population characteristics, and time to seizure

(A) More than 42 000 EEGs were reported using the templated system, most of these being routine (<1 h) EEGs or hospital-based long-term monitoring (CEEG). (B) Completion index for neonatal CEEGs, defined as proportion of key features described, has improved over time from less than 80% when templates were first instituted to more than 95% currently. The dots indicate mean completion indices, and the line shows the trend line fit. (C) All key features were reliably reported for both the overall neonatal cohort and subgroup of neonates with hypoxic ischaemic encephalopathy (solid bars). Frequencies of abnormalities varied slightly between the overall cohort and subgroup of neonates with hypoxic ischaemic encephalopathy (translucent bars). (D) Among the neonates who ultimately had at least one seizure, the duration from CEEG initiation to the first seizure was similar between the overall cohort and subgroup of neonates with hypoxic ischaemic encephalopathy, with most patients having seizures within the first 2 days of monitoring. CEEG=continuous electroencephalogram. EEG=electroencephalogram. ICU=intensive care unit.
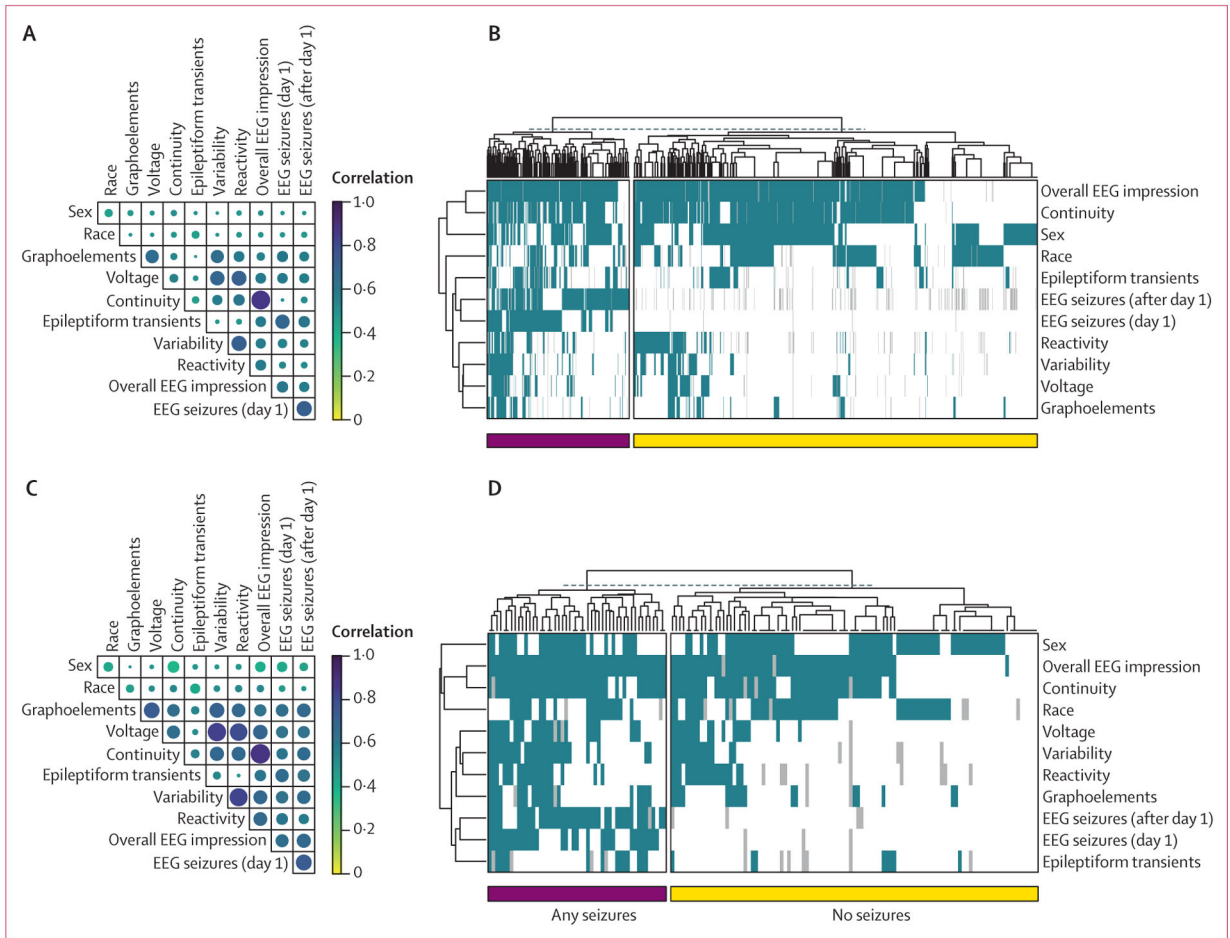
Figure 2: Feature correlation and clustering

Cross-correlation coefficients for all EEG features are shown for the entire neonatal cohort (A) and the subgroup of neonates with hypoxic ischaemic encephalopathy (C). Features with the strongest correlations include impression and continuity. Clustering of individuals on the basis of EEG feature representation did not clearly segment patients based on the presence or absence of seizures for the overall cohort (B) or the subgroup of neonates with hypoxic ischaemic encephalopathy (D). Seizures (day 1) are those occurring on the first day of EEG recording, whereas seizures (after day 1) represent all subsequent seizures. EEG=electroencephalogram.
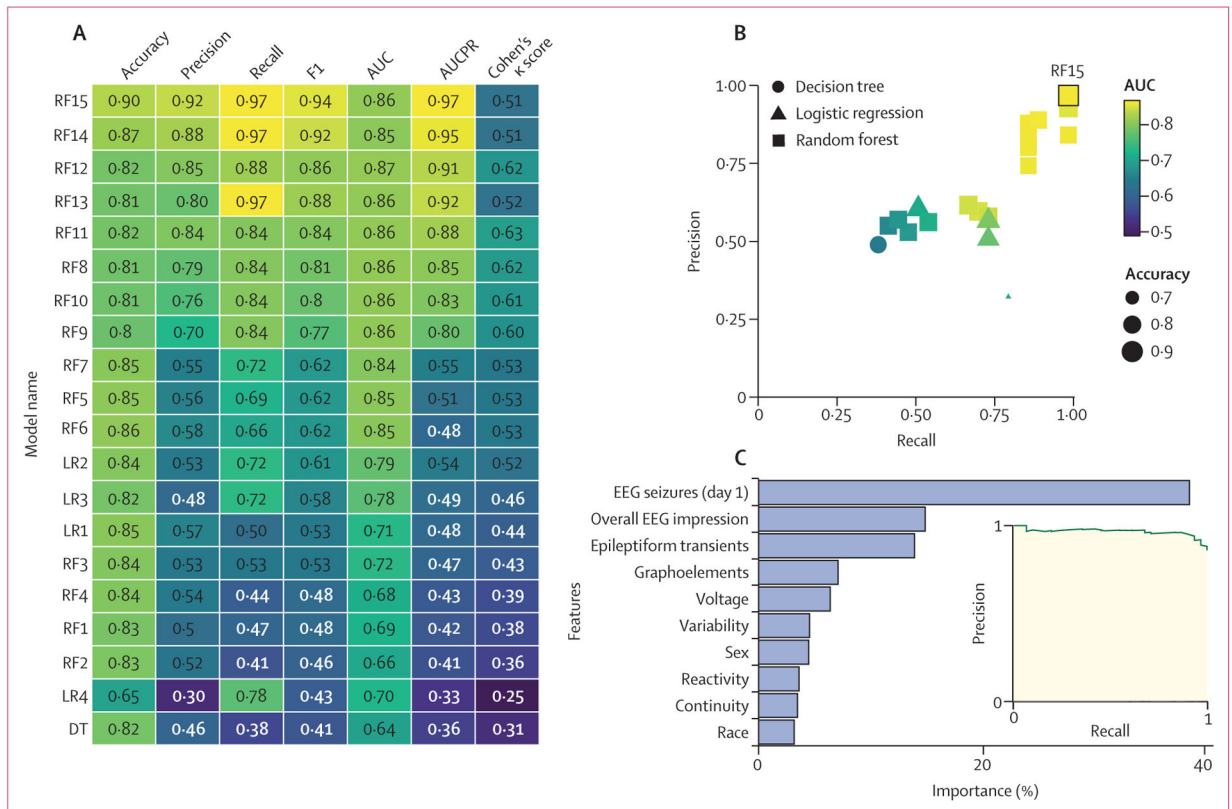
Figure 3: Model performance for all neonates

(A) Performance values are displayed for the logistic regression (LR1–LR4), decision tree (DT), and random forest (RF1–RF15) models tested on the entire neonatal cohort. Accuracy, precision, recall, F1, AUC, AUCPR, and Cohen's κ scores are provided for each model. Lighter colours represent better performance. (B) The precision of each model plotted over recall for all 20 models, coded by type, AUC, and accuracy. (C) The relative importance of each feature in the model and the AUCPR are shown for model RF15. AUC=area under the curve. AUCPR=area under the precision–recall curve.
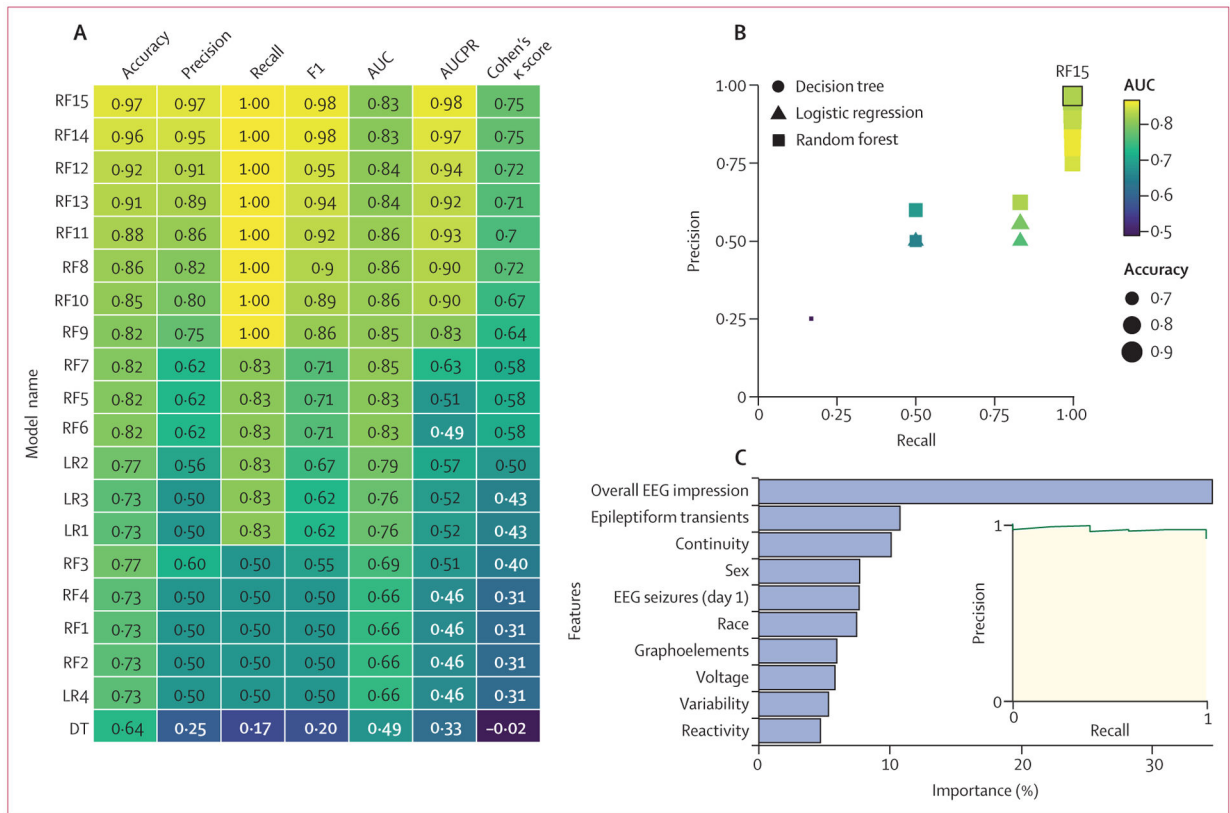
Figure 4: Model performance for neonates with hypoxic ischaemic encephalopathy

Performance values are displayed for the logistic regression (LR1–LR4), decision tree (DT), and random forest (RF1–RF15) models tested on the cohort of neonates with hypoxic ischaemic encephalopathy. Accuracy, precision, recall, F1, AUC, AUCPR, and Cohen's κ scores are provided for each model. Lighter colours represent better performance. (B) The precision of each model plotted over recall for all 20 models, coded by type, AUC, and accuracy. (C) The relative importance of each feature in the model and the AUCPR are shown for model RF15. AUC=area under the curve. AUCPR=area under the precision–recall curve.

Table:

Demographics

	All neonates (n=1117)	Neonates with hypoxic ischaemic encephalopathy (n=150)	p value
Sex			
Male	638 (57%)	92 (61%)	0·33
Female	479 (43%)	58 (39%)	0·33
Race or ethnicity			
White	630 (56%)	75 (50%)	0·16
Black	183 (16%)	34 (23%)	0·064
Asian	40 (4%)	5 (3%)	1
Hispanic or Latinx	175 (16%)	18 (12%)	0·28
Age at CEEG initiation, days	4 (2–7)	1 (1–1)	<0·0001 [*]
CEEG duration, days	3 (2–4)	5 (4–6)	<0·0001 [*]

Data are n (%) or median (IQR) unless otherwise stated. All p values are Fisher's exact test comparing all neonates with a subgroup of neonates with hypoxic ischaemic encephalopathy unless otherwise stated. CEEG=continuous electroencephalogram.

[*] Wilcoxon rank-sum test.