



ELSEVIER

Contents lists available at [ScienceDirect](http://www.sciencedirect.com)

Seizure

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

Development and validation of a seizure prediction model in critically ill children



Amy Yang^k, Daniel H. Arndt^b, Robert A. Berg^m, Jessica L. Carpenter^c, Kevin E. Chapman^d, Dennis J. Dlugos^a, William B. Gallentine^e, Christopher C. Giza^f, Joshua L. Goldstein^g, Cecil D. Hahn^h, Jason T. Lerner^f, Tobias Loddenkemperⁱ, Joyce H. Matsumoto^f, Kendall B. Nash^j, Eric T. Payne^h, Iván Sánchez Fernándezⁱ, Justine Shults^{k,1}, Alexis A. Topjian^m, Korwyn Williams^l, Courtney J. Wusthoffⁿ, Nicholas S. Abend^{a,*}

^a Departments of Neurology and Pediatrics, The Children's Hospital of Philadelphia and The Perelman School of Medicine at The University of Pennsylvania, Philadelphia, PA, United States

^b Departments of Pediatrics and Neurology, Beaumont Children's Hospital and Oakland University William Beaumont School of Medicine, Royal Oak, MI, United States

^c Department of Neurology, Children's National Medical Center, Washington, DC, United States

^d Department of Pediatrics and Neurology, Children's Hospital Colorado and University of Colorado School of Medicine, Aurora, CO, United States

^e Division of Neurology, Duke Children's Hospital and Duke University School of Medicine, Durham, NC, United States

^f Division of Neurology, Department of Pediatrics Mattel Children's Hospital and UCLA Brain Injury Research Center, Department of Neurosurgery, David Geffen School of Medicine at UCLA, Los Angeles, CA, United States

^g Division of Neurology, Children's Memorial Hospital and Northwestern University Feinberg School of Medicine, Chicago, IL, United States

^h Division of Neurology, The Hospital for Sick Children and University of Toronto, Toronto, ON, United States

ⁱ Division of Epilepsy and Clinical Neurophysiology, Department of Neurology, Boston Children's Hospital and Harvard Medical School, Boston, MA, United States

^j Department of Neurology, University of California San Francisco, San Francisco, CA, United States

^k Department of Biostatistics and Epidemiology, Center for Clinical Epidemiology and Biostatistics, Perelman School of Medicine at The University of Pennsylvania, United States

^l Department of Pediatrics, University of Arizona College of Medicine and Barrow's Neurological Institute at Phoenix Children's Hospital, Phoenix, AZ, United States

^m Department of Anesthesia and Critical Care Medicine, The Children's Hospital of Philadelphia and The Perelman School of Medicine at The University of Pennsylvania, Philadelphia, PA, United States

ⁿ Division of Child Neurology, Stanford University, Palo Alto, CA, United States

ARTICLE INFO

Article history:

Received 13 August 2014

Received in revised form 25 September 2014

Accepted 29 September 2014

Keywords:

EEG monitoring

Seizure

Status epilepticus

Pediatric

Non-convulsive seizure

Prediction model

ABSTRACT

Purpose: Electrographic seizures are common in encephalopathic critically ill children, but identification requires continuous EEG monitoring (CEEG). Development of a seizure prediction model would enable more efficient use of limited CEEG resources. We aimed to develop and validate a seizure prediction model for use among encephalopathic critically ill children.

Method: We developed a seizure prediction model using a retrospectively acquired multi-center database of children with acute encephalopathy without an epilepsy diagnosis, who underwent clinically indicated CEEG. We performed model validation using a separate prospectively acquired single center database. Predictor variables were chosen to be readily available to clinicians prior to the onset of CEEG and included: age, etiology category, clinical seizures prior to CEEG, initial EEG background category, and inter-ictal discharge category.

Results: The model has fair to good discrimination ability and overall performance. At the optimal cut-off point in the validation dataset, the model has a sensitivity of 59% and a specificity of 81%. Varied cut-off points could be chosen to optimize sensitivity or specificity depending on available CEEG resources.

* Corresponding author at: Departments of Neurology and Pediatrics, The Children's Hospital of Philadelphia and The Perelman School of Medicine at The University of Pennsylvania, Philadelphia, PA, United States. Tel.: +1 215 590 1719; fax: +1 215 590 1771.

E-mail addresses: abend@chop.edu, abend@email.chop.edu (N.S. Abend).

¹ Dr. Shults is a PhD biostatistician at Children's Hospital of Philadelphia and the Perelman School of Medicine at the University of Pennsylvania.

Conclusion: Despite inherent variability between centers, a model developed using multi-center CEEG data and few readily available variables could guide the use of limited CEEG resources when applied at a single center. Depending on CEEG resources, centers could choose lower cut-off points to maximize identification of all patients with seizures (but with more patients monitored) or higher cut-off points to reduce resource utilization by reducing monitoring of lower risk patients (but with failure to identify some patients with seizures).

© 2014 British Epilepsy Association. Published by Elsevier Ltd. All rights reserved.

1. Introduction

Electrographic seizures (ES) are reported in 10–50% of children with acute encephalopathy who undergo continuous EEG monitoring (CEEG),^{1–19} and there is increasing evidence that high ES burdens are associated with worse outcomes, even in models that adjust for acute encephalopathy etiology and severity.^{10,13,17,20,21} Most ES in critically ill children have no clinical correlate so CEEG is required for identification,^{3,6,8–10,12–15,18,19} leading to recent increases in CEEG use within pediatric intensive care units (PICUs).²² CEEG is resource intense, and seemingly small changes in CEEG utilization may have substantial impacts on equipment and personnel needs.²³ Seizure prediction models could allow CEEG to be targeted to children at highest risk for experiencing ES within the resource limitations of an individual medical center.

Several studies have described clinical and EEG risk factors for ES. However, these data are limited in several ways. First, the data are obtained from patients who have undergone CEEG at single or few institutions,^{3,7–9,11,12,14,17,18} so the risk factors identified may not be useful if implemented at a different institution with different patient characteristics or CEEG practice. Although multi-center consortia are starting to study ES epidemiology,^{13,24,25} multi-center data variability might preclude meaningful application at individual centers. Second, prior studies have not combined the identified risk factors to create clinically useful ES prediction models accounting for multiple risk factors. Currently, a clinician may consider multiple known seizure risk factors (such as younger age, prior convulsive seizures, and inter-ictal epileptiform discharges) when making a clinical judgment regarding the need for CEEG, but there are no data available to help determine the accuracy of this combinatorial approach.

We aimed to determine whether an ES prediction model developed from retrospective multi-center data could be used to predict ES occurrence when applied to data obtained from a single center.

2. Methods

2.1. Datasets

The model was created and validated using separate datasets of children in PICUs who underwent clinically indicated CEEG. The overall study was approved by The Children's Hospital of Philadelphia institutional review board, and submission of data was approved by the institutional review boards at each site.

The model creation dataset was derived from a multi-center study in which 11 sites each collected data by retrospective chart review on 50 consecutive critically ill children to yield 550 subjects.^{13,24,25} The 11 sites were large academic medical centers with available pediatric neurology consultation and CEEG. Subjects had undergone clinically indicated CEEG as dictated by practice patterns at each institution and not any study protocol or national guideline. Thus, these subjects were heterogeneous in terms of etiology, degree of encephalopathy, and other clinical characteristics. For the current study, we excluded 214 subjects with

pre-existing epilepsy-related diagnoses leading to PICU admission, leading to a cohort of 336 subjects. Patients with epilepsy were excluded for several reasons. Prior classification proposals have differentiated between non-convulsive status epilepticus (NCSE) as occurring in the context of acute brain injury (termed “comatose NCSE”) and occurring in the context of more benign epilepsy conditions (termed “NCSE proper”) since the relative impact of seizures to overall prognosis differs.²⁶ Our aim was to address the use of CEEG when screening for electrographic seizures in patients with acute encephalopathy in whom seizure identification and management might serve as a neuroprotective strategy. Second and more practically, institutional practice, bed availability, and admission time of day likely impact decisions regarding whether patients with epilepsy in need of CEEG are admitted to the epilepsy monitoring unit or PICU. Data were obtained by chart review from the reports created by trained encephalographers on-service when the CEEG was obtained and the tracings were not re-interpreted for this study.

The model validation dataset was derived from a prospective single center dataset from The Children's Hospital of Philadelphia and included 222 subjects who underwent CEEG while in the PICU without an epilepsy-related diagnosis prior to admission. As described above, patients with epilepsy were excluded. Institutional practice at The Children's Hospital of Philadelphia is to obtain at least one day of CEEG in any patient admitted to the PICU with encephalopathy of any degree and any acute neurologic condition (i.e. traumatic brain injury, stroke, hypoxic ischemic encephalopathy, encephalitis). These were different subjects than those contributed by The Children's Hospital of Philadelphia to the multi-center model creation dataset. Epidemiologic data from a portion of this dataset have been published previously.¹⁸ The single center validation data were obtained prospectively by one investigator (N.S.A.) who re-scored the CEEG after clinical interpretation for seizure category and background category while blind to clinical data other than age.

We categorized subjects by ES category (none, electrographic seizures, electrographic status epilepticus). Electrographic status epilepticus was defined as a single or recurrent electrographic seizure(s) lasting 30 min or more within a 1 hour epoch. We collected data regarding clinical variables previously identified as predicting an increased risk of experiencing ES. Age was classified as >24 months or ≤24 months. Clinically evident seizures prior to CEEG were classified as present or absent. Etiology category was classified as structural (i.e. traumatic brain injury, stroke (ischemic or hemorrhagic), hypoxic-ischemic encephalopathy, encephalitis, posterior reversible leukoencephalopathy) or non-structural (hepatic encephalopathy, sepsis, metabolic disorders). The initial 1 h EEG background categorization (normal/sleep, slow-disorganized, discontinuous, burst-suppression, or attenuated/featureless), and initial one hour EEG inter-ictal epileptiform discharge (IED) categorization (present or absent). This background categorization scheme has been used in prior studies related to EEG in critically ill children.^{13,20,21} When EEGs from critically ill children are reviewed by pediatric encephalographers, inter-rater agreement for background features continuity (continuous, discontinuous, flat) and

Table 1
Summary characteristics of the creation and validation datasets.

| | Creation dataset | Validation dataset | p-Value |
|---|----------------------|-----------------------|---------|
| Cohort | Multi-center (N=336) | Single-center (N=222) | |
| Male N (%) | 180 (53.57) | 130 (58.56) | 0.283 |
| Age | | | |
| Median (IQR) | 28.4 (6–122.2) | 32.37 (7.13–125.8) | |
| <2 years N (%) | 161 (47.92) | 97 (43.69) | |
| Clinically evident seizures prior to CEEG N (%) | 142 (42.26) | 99 (44.59) | 0.648 |
| Typical EEG background category N (%) | | | 0.381 |
| Normal/sleep | 67 (19.94) | 41 (18.47) | |
| Slow/disorganized | 182 (54.17) | 138 (62.16) | |
| Discontinuous | 26 (7.74) | 12 (5.41) | |
| Burst-suppression | 22 (6.55) | 12 (5.41) | |
| Attenuated/featureless | 39 (11.61) | 19 (8.56) | |
| Inter-ictal epileptiform discharges N (%) | 107 (31.85) | 50 (22.52) | 0.021 |
| Etiology N (%) | | | |
| Acute structural | 212 (63.1) | 168 (75.68) | 0.0025 |
| Acute non-structural | 124 (36.9) | 54 (24.32) | |
| Electrographic seizures N (%) | 70 (20.83) | 83 (37.39) | <0.001 |

CEEG, continuous EEG monitoring; IQR, interquartile range.

burst suppression (present or absent) have been substantial and agreement for inter-ictal discharges (present or absent) has been fair.²⁷

2.2. Analysis²

Descriptive statistics were calculated for the model creation and model validation datasets, and the Pearson's χ^2 statistic was used to compare differences.

Predictive variables were first analyzed by univariate logistic regressions with statistical significance denoted by a two-sided p-value of <0.05. Significant variables were entered into a multivariate logistic regression model and the backward selection method was used to generate the final model. Individual predictors' ability to classify seizure categorization was assessed using beta coefficients and their corresponding z-statistics and p-values.

Overall performance was measured by explained variation (Nagelkerke's R^2)²⁸ and Brier score/Brier_{scaled}. R^2 ranges from 0 to 1, and higher values indicate more variation is explained by the predictor. The Brier score ranges from 0 (perfect model) to 0.25 (non-informative model). We scaled the Brier score by its maximum score under a non-informative model: Brier_{scaled} = $1 - \text{Brier}/\text{Brier}_{\text{max}}$, where Brier_{max} = $\text{mean}(p) \times (1 - \text{mean}(p))$, to let it range between 0% and 100% to ease the interpretation.

Accurate predictions discriminate between subjects with and without ES. The concordance statistic (c-statistic) was used to indicate the discriminative ability of the logistic regression model. The c-statistic represents the probability that a subject who experiences an ES has a higher predicted probability than a seizure absent subject for a random pair of subjects consisting of one subject with and one subject without an ES. It is equivalent to the area under the receiver operating characteristic (ROC) curve (AUC) for binary outcomes.²⁹ Higher values were considered to demonstrate better discrimination abilities as follows: excellent (AUC ≥ 0.90), good ($0.80 \leq \text{AUC} < 0.90$), fair ($0.70 \leq \text{AUC} < 0.80$), and poor (AUC < 0.70).³⁰ In addition, the discrimination slope was used to measure how well subjects with and without ES were separated. It was calculated as the absolute difference in average predictions for subjects with and without ES.

The validity of the predicted SE was measured graphically by plotting the predicted outcomes on the x-axis and the observed outcomes on the y-axis. This calibration plot assesses the agreement between the observed outcomes and predicted

outcomes. A 45-degree line would represent perfect prediction. Since the outcomes in the study were binary with only 0 or 1 value for the y-axis, we plotted observed outcome by decile of predictions and added smoothing line by LOESS (locally estimated scatterplot smoothing). The closeness of the smoothing line and the 45-degree line is a visualization of the agreement/calibration.³¹ The performance for calibration was measured using a recalibration model $\text{logit}(y) = a + b \times \text{linear predictors}$ proposed by Cox.³² The values for a and b are two indicators for calibration performance. For regression model development, $a = 0$ and $b = 1$ were set as fixed. For validation, a departure from zero for intercept a indicated that the predictions were systematically too low or too high given that b set to 1. A value of calibration slope b smaller than 1 given that a set to 0 reflected over-fitting of the model, indicating that the predictions are too extreme (overly low estimates for low predictions and overly high estimates for high predictions). The observed outcome by decile of prediction in this calibration plot is equivalent to a graphical illustration of the Hosmer–Lemeshow test,³³ which was used to test agreement between the observed frequencies and predicted probabilities. A high Hosmer–Lemeshow statistic is related to a small p-value and indicates lack of fit.

Jack-knife cross-validation was used on the creation dataset as a measure of internal validation to obtain an overall prediction error. The model was then applied to the external dataset to test its performance using all the measures above.

The predicted ES probabilities were calculated from the model, and they were termed *model scores*. Sensitivity and specificity were obtained from various model score cut-offs, assuming patients receiving model scores above the cut-off would be selected to undergo CEEG. Positive predictive values (PPV) and negative predictive values (NPV) were calculated using a 30% ES prevalence based on multi-center seizure prevalence data.¹³

3. Results

3.1. Descriptive statistics

Table 1 provides summary characteristics regarding the clinical and EEG variables from the model creation and model validation datasets. The two datasets were similar in the percentages of males, subjects younger than two years, occurrence of clinically evident seizures prior to CEEG, and typical background EEG category. The model creation dataset contained a higher proportion of subjects with IEDs and the model validation dataset

² Amy Yang and Dr. Justine Shults performed the statistical analyses.

contained a higher proportion of subjects with ES and with structural neurologic diagnoses.

3.2. Model creation dataset

The best model for discriminating between ES category outcomes was identified using logistic regression. All variables used in the univariate analysis remained significant in the multivariate analysis except for Etiology Category (Wald $z = -1.47, p = 0.1423$), which was removed from the final model. Table 2 provides the beta-coefficients and odds ratios from this model. The strongest predictor of ES presence was EEG background; seizures were more likely in subjects with burst-suppression and attenuated/featureless backgrounds. Other

variables associated with higher odds of ES occurrence were: age ≤ 24 months, clinically evident seizures prior to CEEG, and the presence of IEDs. Table 3 provides the final regression formula as well as the calculated risk values (model scores) for each possible combination of variables. Supplemental Fig. 1 provides a graphical summary of the model scores.

A model with good discrimination will show a wide spread in the distribution of the predicted probabilities away from the average probability. Fig. 1 shows the distribution of the probabilities for subjects per outcome value (seizures absent or present). Subjects who experienced ES had 27% higher predicted ES probabilities than subjects who did not experience ES. The c -statistic/area under the curve was 0.845 (Fig. 1) indicating the model had good discrimination ability in the model creation dataset.

Table 2
Risk factors in logistic regression model in model creation dataset.

| Variable | Log OR (beta) | SE (beta) | z | p-Value | OR | OR 95% CI |
|---|---------------|-----------|------|---------|-------|-------------|
| Age (<24 months) | 0.91 | 0.33 | 2.76 | 0.0058 | 2.48 | 1.30–4.72 |
| Clinically evident seizures prior to CEEG (present) | 0.81 | 0.32 | 2.49 | 0.0126 | 2.24 | 1.19–4.22 |
| Typical EEG background category (ref=normal/sleep) | | | | | | |
| Slow/disorganized | 2.43 | 1.04 | 2.34 | 0.0192 | 11.40 | 1.49–87.39 |
| Discontinuous | 2.22 | 1.13 | 1.97 | 0.0488 | 9.24 | 1.01–84.4 |
| Burst-suppression | 3.08 | 1.15 | 2.69 | 0.0072 | 21.79 | 2.31–205.91 |
| Attenuated/featureless | 2.97 | 1.10 | 2.7 | 0.0069 | 19.42 | 2.25–167.33 |
| Inter-ictal epileptiform discharges (present) | 2.43 | 0.40 | 6.14 | <0.0001 | 11.34 | 5.22–24.62 |

CEEG, continuous EEG monitoring; CI, confidence interval; OR, odds ratio.

Table 3
The final logistic regression formula for a subject's risk of electrographic seizure occurrence based on the above values is given by the formula: $\text{risk} = e^w / (1 + e^w)$ where the linear predictor $w = 0.9076 \cdot \text{age}$ (1 = less than 24 months, 0 = greater or equal to 24 months) + $0.8056 \cdot \text{clinically evident seizures prior to CEEG}$ (1 = yes, 0 = no) + $2.4334 \cdot \text{slow and disorganized initial background}$ (1 = yes; 0 = no) + $2.2239 \cdot \text{discontinuous initial background}$ (1 = yes; 0 = no) + $3.0813 \cdot \text{burst-suppression initial background}$ (1 = yes; 0 = no) + $2.9661 \cdot \text{attenuated/featureless initial background}$ (1 = yes; 0 = no) + $1.8272 \cdot \text{interictal epileptiform discharges in the first hour of EEG}$ (1 = present, 0 = absent) – 4.5067. To avoid requiring the calculations in clinical use, the model score is calculated for each combination of variables using the logistic regression formula. Patients with model scores above an institution's pre-determined cut-off would undergo CEEG.

| Age | Clinically evident seizures prior to CEEG | EEG background category | Interictal epileptiform discharges | Model score |
|------------|---|----------------------------|------------------------------------|-------------|
| <24 months | Yes | Slow and disorganized | Yes | 0.8126 |
| <24 months | Yes | Discontinuous | Yes | 0.7786 |
| <24 months | Yes | Burst suppression | Yes | 0.8924 |
| <24 months | Yes | Attenuated and featureless | Yes | 0.8808 |
| <24 months | Yes | Slow and disorganized | No | 0.4110 |
| <24 months | Yes | Discontinuous | No | 0.3613 |
| <24 months | Yes | Burst suppression | No | 0.5715 |
| <24 months | Yes | Attenuated and featureless | No | 0.5430 |
| <24 months | No | Slow and disorganized | Yes | 0.6596 |
| <24 months | No | Discontinuous | Yes | 0.6111 |
| <24 months | No | Burst suppression | Yes | 0.7874 |
| <24 months | No | Attenuated and featureless | Yes | 0.7675 |
| <24 months | No | Slow and disorganized | No | 0.2376 |
| <24 months | No | Discontinuous | No | 0.2018 |
| <24 months | No | Burst suppression | No | 0.3734 |
| <24 months | No | Attenuated and featureless | No | 0.3468 |
| ≥24 months | Yes | Slow and disorganized | Yes | 0.6363 |
| ≥24 months | Yes | Discontinuous | Yes | 0.5866 |
| ≥24 months | Yes | Burst suppression | Yes | 0.7698 |
| ≥24 months | Yes | Attenuated and featureless | Yes | 0.7488 |
| ≥24 months | Yes | Slow and disorganized | No | 0.2197 |
| ≥24 months | Yes | Discontinuous | No | 0.1859 |
| ≥24 months | Yes | Burst suppression | No | 0.3498 |
| ≥24 months | Yes | Attenuated and featureless | No | 0.3241 |
| ≥24 months | No | Slow and disorganized | Yes | 0.4388 |
| ≥24 months | No | Discontinuous | Yes | 0.3880 |
| ≥24 months | No | Burst suppression | Yes | 0.5991 |
| ≥24 months | No | Attenuated and featureless | Yes | 0.5712 |
| ≥24 months | No | Slow and disorganized | No | 0.1117 |
| ≥24 months | No | Discontinuous | No | 0.0926 |
| ≥24 months | No | Burst suppression | No | 0.1938 |
| ≥24 months | No | Attenuated and featureless | No | 0.1764 |

CEEG, continuous EEG monitoring.

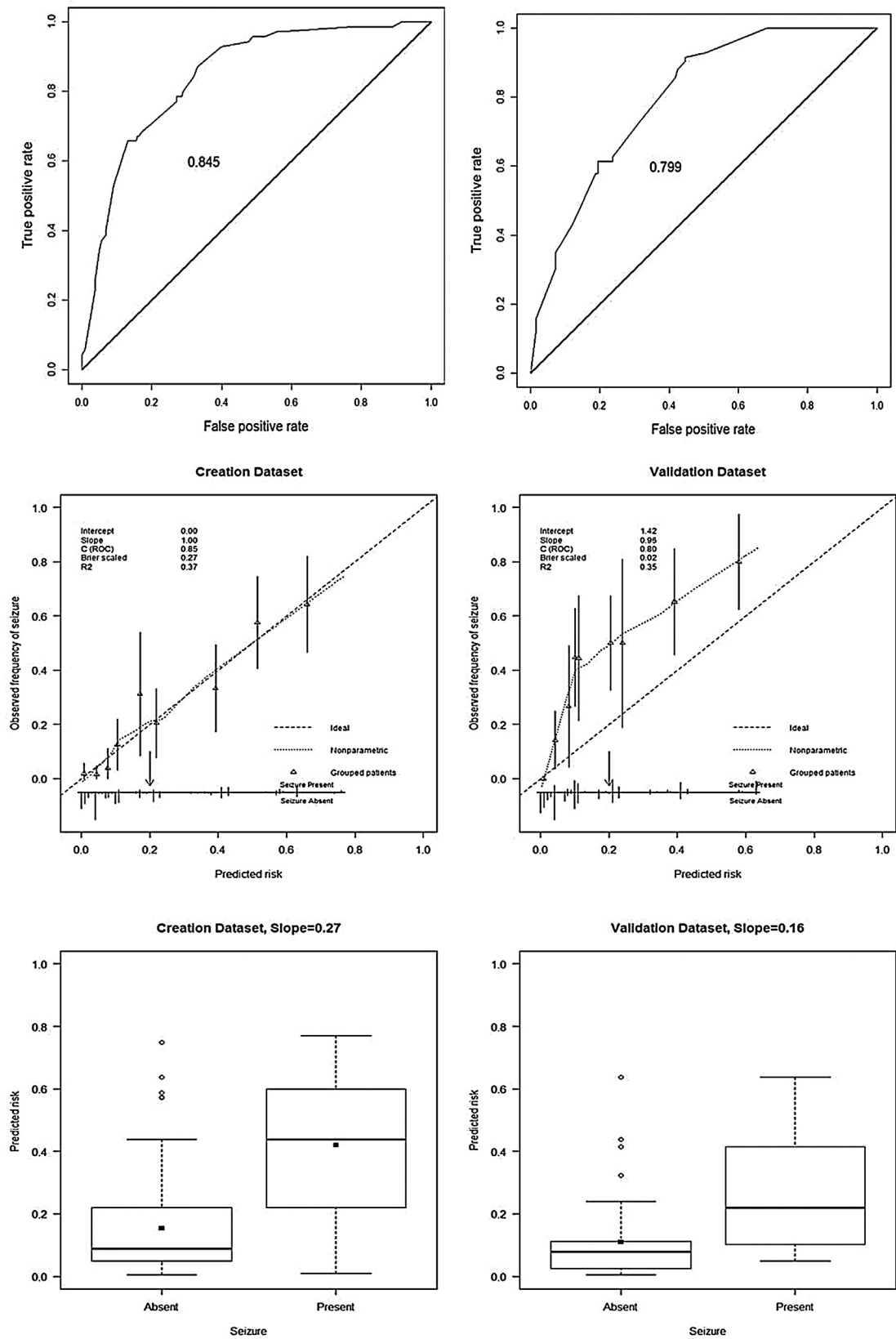


Fig. 1. Characteristics are provided for the creation (left) and validation (right) datasets. *Top.* Receiver operating characteristic (ROC) curves. *Middle.* Validation plots of prediction model for residual masses in creation data and validation data. Calibration is shown with a non-parametric line (dot). Distribution of the predicted probabilities and 95% confidence intervals are indicated by groups of patients (triangles with vertical line) and for individual patients (vertical lines). Vertical lines upwards represent patients with seizures; lines downwards represent patients without seizures. The threshold value was chosen to be 20% (downwards arrow). *Bottom.* Box plots of predicted probabilities in creation and validation datasets. Group means are displayed as filled squares. Points outside 1.5 interquartile ranges are displayed as unfilled circles.

Table 4

Using the logistic regression formula or the pre-calculated resulting model scores for each combination of variables shown in Table 3 a patient's model score would be identified. Patients with model scores above an institutionally determined model score cut-off would undergo CEEG. For various cut-offs the sensitivity and specificity are provided for the creation and validation datasets. Compared to all patients who underwent CEEG in the datasets, the percentage of patients who would be above the model score cutoff and thus undergo CEEG are provided for each cutoff. Bold font indicates the model score cut-off maximizing sensitivity and specificity.

| Model score cut-off | | 0.10 | 0.15 | 0.20 | 0.25 | 0.35 | 0.45 |
|---------------------|---|-------|-------|--------------|-------|-------|-------|
| Creation dataset | Sensitivity/specificity | 94/52 | 87/67 | 79/73 | 72/79 | 67/84 | 53/91 |
| | Percentage of patients classified as needing CEEG at each cut-off | 36 | 28 | 24 | 18 | 17 | 9 |
| Validation dataset | Sensitivity/specificity | 86/58 | 62/76 | 59/81 | 43/88 | 34/92 | 19/97 |
| | Percentage of patients classified as needing CEEG at each cut-off | 43 | 28 | 25 | 14 | 13 | 5 |

Table 4 provides the sensitivities and specificities using different model scores cut-offs. In practice, patients with model scores above the cut-off would be considered at higher risk for experiencing ES and would undergo CEEG. A model score cut-off of 0.2 maximized sensitivity and specificity in the creation model. The model worked well in the creation dataset with a Brier score of 0.121 indicating an informative prediction. The Hosmer–Lemeshow goodness-of-fit test was not significant ($\chi^2 = 6.01$, p -value = 0.65) indicating the model had adequate fit for this dataset. In the model creation dataset, the prediction error from jack-knife cross-validation was 13%, suggesting that approximately 87% of the cases would be correctly classified.

3.3. Model validation dataset

Overall model performance in the model validation dataset was less than in the creation dataset, according to R^2 (35% for validation dataset in contrast to 37% for creation dataset) and Brier score (0.23 for validation dataset in contrast to 0.12 for creation dataset). The c -statistic AUC was 0.799, indicating fair discrimination ability in the model validation dataset. The discrimination slope was 0.16 (Fig. 1) indicating on average the difference in the predicted probability between the two outcome groups was 16%.

Fig. 1 provides the calibration plots of the model predicting ES presence for the model validation dataset. The intercept and slope of the calibration line were estimated in a logistic regression model with the linear predictor w calculated for the patients of the validation dataset, as the only predictor variable: $\log \text{odds} = a + b$ linear predictor. The calibration intercept was 1.42 which is greater than zero indicating the predicted probabilities were systematically too low. The slope was 0.96 which is close to 1 and therefore indicates that the model is not too optimistic (over-fitting). This type of miscalibration (slope close to 1 and the intercept different from 0) indicates that certain patient characteristics not included in the prediction model were differently distributed in the validation sample compared with the model creation sample. The Hosmer–Lemeshow test was statistically significant ($\chi^2 = 104$, p -value < 0.0001) indicating a lack of fit in the model validation dataset, which agrees with the calibration plot.

The sensitivities and specificities using different model score cut-offs are shown in Table 4. Using the optimal 0.2 model score cut-off identified in the model creation dataset, only 25% of patients in the model validation dataset who might have undergone clinically indicated CEEG would be selected to undergo CEEG. This practice would lead to lower resource requirements but would fail to identify some patients experiencing ES.

4. Discussion

A seizure prediction model derived from a retrospectively acquired multi-center dataset provides useful ES predictions when validated using a single center dataset. Having a model that discriminates well between patients who will and will not experience ES can improve resource utilization since limited CEEG

resources can be directed to the patients most at risk for experiencing ES. While imperfect, our model has a fair to good discrimination even on the validation dataset indicating that most patients could be appropriately classified.

Many studies have indicated that 10–50% of children with acute encephalopathy who undergo CEEG experience ES.^{1–19} Because CEEG is resource intense and a limited number of patients might be able to undergo CEEG at a given center, using this model derived from a small number of clinically evident risk factors might improve CEEG efficiency. This model could be applied clinically in a few steps. First, the clinician must know several clinical variables and EEG variables which could be obtained from a routine EEG. Second, using these variables, the clinician can calculate the model score using the logistic regression formula or, more easily, by locating the appropriate row in Table 3 or location on Supplemental Fig. 1, both of which provide pre-calculated model scores for each combination of variables. Third, patients with model scores above the institutional cut-off would be selected to undergo CEEG.

Individual institutions might pick different model score cut-offs based on center-specific criteria. A model score cut-off of 0.20 had the best overall test characteristics. Given a seizure prevalence of 30%, the estimated PPV was 57% and NPV was 82%. Only 25% of patients who might have undergone clinically indicated CEEG would be classified as needing CEEG, leading to lower resource requirements. However, 41% of patients experiencing ES would not be classified as needing CEEG and thus the ES would not be identified. A center with substantial CEEG resources might perform CEEG for any patient with a model score >0.10. At this cut-off, 14% of patients with ES would not undergo CEEG so the seizures would not be identified and managed. However, 58% of patients without ES would be identified as not needing CEEG, so limited CEEG resources would not be expended. At this cut-off, only 43% of all patients who might have undergone CEEG would be selected to undergo CEEG. Given a seizure prevalence of 30%, this would result in a PPV of 47% and NPV of 91%. In contrast, a center with more limited CEEG resources might only perform CEEG for any patient with a model score >0.45. At this cut-off, 81% of patients with ES would not undergo CEEG so the seizures would not be identified and managed. However, 97% of patients without ES would be identified as not needing CEEG, so limited CEEG resources would not be expended. Given a seizure prevalence of 30%, this would result in a PPV of 74% and NPV of 74%.

There are limitations to these data. First, the model included a limited number of categorized variables. While this kept the model simple and focused on data readily available to clinicians at admission, utilizing additional clinical and EEG variables might yield better predictive models. For example, etiology was categorized as structural or non-structural, and future larger studies using more exact etiologic categories might yield better predictive models. Additionally, a limited number of clinical variables were included, and future studies might include variables related to critical illness severity, neuroimaging, hypoxia, hypotension, vasopressor use, and markers of injury to other organs. Second, the multi-center dataset used for

validation was generated from 11 large pediatric centers based on clinically obtained CEEG with retrospective data acquisition. Each institution had different CEEG practices and PICU referral patterns that lead to variability as well as selection and referral bias for CEEG inclusion. Seizure occurrence was significantly lower in the multi-center dataset than the Children's Hospital of Philadelphia dataset reflecting these cross-institutional differences in CEEG practice and PICU referral patterns. This variability likely underlies the imperfect fit for the single center validation dataset. The lack of optimal model fit cautions clinical investigators aiming to develop evidence-based CEEG pathways that single center studies based on clinically indicated CEEG may have a limited ability to provide generalizable data. Future studies of consecutive patients with very specific CEEG inclusion criteria may be needed to generate improved seizure prediction models. However, the fact that model fit was fair despite this variability and the limited number of variables utilized indicates that with further development data acquired from multi-center studies could have meaningful applicability at individual institutions. Third, the model may under-predict seizures. Developing a model which is well calibrated across diverse settings is difficult since models with a limited number of predictors always leave some variance between subjects unexplained. The intercept from the validation plot was greater than zero and the fitted line between observed seizure frequencies and predicted frequencies was above the 45° diagonal line indicating the model under-predicted seizures on average in the validation dataset. Seizure occurrence was significantly higher in the validation than creation dataset, whereas three out of four predictor variables were distributed similarly between the two datasets. A model developed from a low outcome rate population applied to a higher outcome rate population will produce under-prediction. These data suggest we should aim to identify and include additional predictive variables. Fourth, the model requires information obtained from a routine EEG (EEG background category and inter-ictal epileptiform discharge category) so implementation still requires EEG technologist work. However, selecting patients who need only routine EEGs and not CEEG may still be useful by reducing technologist work time at bedside (electrode placement may be quicker for many patients since electrodes could be applied with paste rather than collodion), physician work time (fewer patients undergoing CEEG would reduce the pages of EEG to be reviewed each day) and equipment needs (collector equipment would not need to remain at bedside for CEEG). Fifth, ES occurrence and background category were determined by a number of encephalographers from each site for the multi-center dataset and one encephalographer for the single center dataset. Further, the multi-center data was obtained by chart review of CEEG reports and not by re-review of the CEEG tracings. There were likely some inter-rater differences in EEG interpretation.²⁷ Multi-rater scoring using unified seizure definitions^{34,35} may improve model performance.

We created the prediction model using multi-center data and validated it using single center data. Studying ES epidemiology and impact on outcome will likely require multi-center studies to enroll sufficient patients, particularly if patients are to be stratified by injury type. Consortia are already performing initial multi-center studies working toward these goals.^{13,24,25} Our field will need to determine whether these multi-center data can be applied meaningfully at individual institutions despite inherent inter-institutional variability. These considerations guided our decision to develop the model using multi-center data and validate it using single center data, rather than vice versa. The fact that the model had predictive ability indicates that future studies using larger multi-center datasets incorporating more variables could have

meaningful applicability when applied at individual centers. We are working to develop a free app which could easily calculate model scores when predictor variables are entered (Supplementary Fig. 2). Development of this type of tool could make complex predictor models developed over time easily implementable by clinicians.

5. Conclusions

The current model was based on small number of clinically known variables, was developed in a broad cohort of patients who underwent clinically indicated CEEG at 11 institutions and was validated in a broad cohort of patients at a single center. Despite these limitations, the model had fair to good discrimination ability and may help clinicians consider multiple risk factors together when making decisions regarding CEEG use. With further development including more defined subject cohorts and additional predictor variables, these data suggest that prediction models derived from multi-center data could yield improved CEEG efficiency at individual centers.

Conflict of interest statement

This study was performed by the Pediatric Critical Care EEG Group (PCEEG) which is the pediatric subgroup of the Critical Care EEG Monitoring Research Consortium (CEMRC). Dr. Abend has received support from NIH K23NS076550, has been paid as an expert in medico-legal cases, and has been paid royalties from Demos Medical Publishing for *Pediatric Neurocritical Care*. Dr. Dlugos has received support from NIH grants 1R01NS053998, 2U01NS045911, 1R01LM011124, and U01NS077276, and he has been paid as an expert in medico-legal cases. Dr. Giza has received support from the California State Athletic Commission, the Sarah Jane Brain Project, the National Hockey League Players' Association, National Football League, the Major League Soccer, the NCAA, the Advisory Board for the American Association for Multi-Sensory Environments (AAMSE), the NINDS/NIH, University of California, Thrasher Research Foundation, Today's and Tomorrow's Children Fund and the Child Neurology Foundation/Winokur Family Foundation, received royalties from Blackwell Publishing for "Neurological Differential Diagnosis," and has been paid as an expert in medico-legal cases. Dr. Loddenkemper has received support from the Laboratory Accreditation Board for Long Term (Epilepsy and Intensive Care Unit) Monitoring, on the Council of the American Clinical Neurophysiology Society, and the American Board of Clinical Neurophysiology, the National Institutes of Health/NINDS, Harvard Medical School and Boston Children's Hospital, the Payer Provider Quality Initiative, The American Epilepsy Society, The Epilepsy Foundation of America, the Center for Integration of Medicine and Innovative Technology, the Epilepsy Therapy Project, the Pediatric Epilepsy Research Foundation, the Danny Did Foundation, the HHV6 Foundation, and from investigator initiated research grants from Lundbeck and Eisai. Dr. Hahn has received support from the Council of the American Clinical Neurophysiology Society, the Canadian Society of Clinical Neurophysiologists, the Canadian Association of Child Neurology, the Canadian Institutes of Health Research, the SickKids Foundation, the PSI Foundation and has been paid as an expert in medico-legal cases. Dr. Sánchez Fernández has received support from the Fundación Alfonso Martín Escudero and the HHV6 Foundation. Dr. Payne has received support from Alberta Innovates Health Solutions and the Canadian League Against Epilepsy. The remaining authors have no conflicts of interest.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.seizure.2014.09.013>.

References

- Abend NS, Wusthoff CJ, Goldberg EM, Dlugos DJ. Electrographic seizures and status epilepticus in critically ill children and neonates with encephalopathy. *Lancet Neurol* 2013;12:1170–9.
- Hosain SA, Solomon GE, Kobylarz EJ. Electroencephalographic patterns in unresponsive pediatric patients. *Pediatr Neurol* 2005;32:162–5.
- Jette N, Claassen J, Emerson RG, Hirsch LJ. Frequency and predictors of non-convulsive seizures during continuous electroencephalographic monitoring in critically ill children. *Arch Neurol* 2006;63:1750–5.
- Abend NS, Dlugos DJ. Nonconvulsive status epilepticus in a pediatric intensive care unit. *Pediatr Neurol* 2007;37:165–70.
- Tay SK, Hirsch LJ, Leary L, Jette N, Wittman J, Akman CI. Nonconvulsive status epilepticus in children: clinical and EEG characteristics. *Epilepsia* 2006;47:1504–9.
- Shahwan A, Bailey C, Shekerdemian L, Harvey AS. The prevalence of seizures in comatose children in the pediatric intensive care unit: a prospective video-EEG study. *Epilepsia* 2010;51:1198–204.
- Abend NS, Topjian AA, Ichord R, Herman ST, Helfaer M, Donnelly M, et al. Electroencephalographic monitoring during hypothermia after pediatric cardiac arrest. *Neurology* 2009;72:1931–40.
- Williams K, Jarrar R, Buchhalter J. Continuous video-EEG monitoring in pediatric intensive care units. *Epilepsia* 2011;52:1130–6.
- Greiner HM, Holland K, Leach JL, Horn PS, Hershey AD, Rose DF. Nonconvulsive status epilepticus: the encephalopathic pediatric patient. *Pediatrics* 2012;129:e748–55.
- Kirkham FJ, Wade AM, McElduff F, Boyd SG, Tasker RC, Edwards M, et al. Seizures in 204 comatose children: incidence and outcome. *Intens Care Med* 2012;28:853–62.
- Arango JL, Deibert CP, Brown D, Bell M, Dvorchik I, Adelson PD. Posttraumatic seizures in children with severe traumatic brain injury. *Child's Nerv Syst* 2012;28:1925–9.
- Piantino JA, Wainwright MS, Grimason M, Smith CM, Hussain E, Byron D, et al. Nonconvulsive seizures are common in children treated with extracorporeal cardiac life support. *Pediatr Crit Care Med* 2013;14:601–9.
- Abend NS, Arndt DH, Carpenter JL, Chapman KE, Cornett KM, Gallentine WB, et al. Electrographic seizures in pediatric ICU patients: cohort study of risk factors and mortality. *Neurology* 2013;81:383–91.
- McCoy B, Sharma R, Ochi A, Go C, Otsubo H, Hutchison JS, et al. Predictors of nonconvulsive seizures among critically ill children. *Epilepsia* 2011;52:1973–8.
- Schreiber JM, Zelleke T, Gaillard WD, Kaulas H, Dean N, Carpenter JL. Continuous video EEG for patients with acute encephalopathy in a pediatric intensive care unit. *Neurocrit Care* 2012;17:31–8.
- Arndt DH, Lerner JT, Matsumoto JH, Madikians A, Yudovin S, Valino H, et al. Subclinical early posttraumatic seizures detected by continuous EEG monitoring in a consecutive pediatric cohort. *Epilepsia* 2013;54:1780–8.
- Payne ET, Zhao XY, Frndova H, McBain K, Sharma R, Hutchison JS, et al. Seizure burden is independently associated with short term outcome in critically ill children. *Brain* 2014;137:1429–38.
- Abend NS, Gutierrez-Colina AM, Topjian AA, Zhao H, Guo R, Donnelly M, et al. Nonconvulsive seizures are common in critically ill children. *Neurology* 2011;76:1071–7.
- Gold JJ, Crawford JR, Glaser C, Sherif H, Wang S, Nespeca M. The role of continuous electroencephalography in childhood encephalitis. *Pediatr Neurol* 2014;50:318–23.
- Topjian AA, Gutierrez-Colina AM, Sanchez SM, Berg RA, Friess SH, Dlugos DJ, et al. Electrographic status epilepticus is associated with mortality and worse short-term outcome in critically ill children. *Crit Care Med* 2013;31:215–23.
- Wagenman KL, Blake TP, Sanchez SM, Schultheis MT, Radcliffe J, Berg RA, et al. Electrographic status epilepticus and long-term outcome in critically ill children. *Neurology* 2014;82(5):396–404.
- Sanchez SM, Carpenter J, Chapman KE, Dlugos DJ, Gallentine W, Giza CC, et al. Pediatric ICU EEG monitoring: current resources and practice in the United States and Canada. *J Clin Neurophysiol* 2013;30:156–60.
- Gutierrez-Colina AM, Topjian AA, Dlugos DJ, Abend NS. EEG monitoring in critically ill children: indications and strategies. *Pediatr Neurol* 2012;46:158–61.
- Sanchez SM, Arndt DH, Carpenter JL, Chapman KE, Cornett KM, Dlugos DJ, et al. Electroencephalography monitoring in critically ill children: current practice and implications for future study design. *Epilepsia* 2013;54:1419–27.
- Sanchez Fernandez I, Abend NS, Arndt DH, Carpenter JL, Chapman KE, Cornett KM, et al. Electrographic seizures after convulsive status epilepticus in children and young adults. A retrospective multicenter study. *J Pediatr* 2014;164:339–46.
- Bauer G, Trinka E. Nonconvulsive status epilepticus and coma. *Epilepsia* 2009;51(2):177–90.
- Abend NS, Gutierrez-Colina AM, Zhao H, Guo R, Marsh E, Clancy RR, et al. Interobserver reproducibility of electroencephalogram interpretation in critically ill children. *J Clin Neurophysiol* 2011;28:15–9.
- Nagelkerke NJD. A note on a general definition of the coefficient of determination. *Biometrika* 1991;78:691–2.
- Hanley JA. Receiver operating characteristic (ROC) methodology: the state of the art. *Crit Rev Diagn Imaging* 1989;29:307–35.
- Fan J, Upadhye S, Worster A. Understanding receiver operating characteristic (ROC) curves. *CJEM* 2006;8:19–20.
- Harrell FE. *Regression modeling strategies*. New York, NY: Springer; 2002.
- Cox DR. Two further applications of a model for binary regression. *Biometrika* 1958;45:562–5.
- Hosmer DW, Lemeshow S. *Applied logistic regression*. Hoboken, NJ: Wiley; 1989.
- Beniczky S, Hirsch LJ, Kaplan PW, Pressler R, Bauer G, Aurlien H, et al. Unified EEG terminology and criteria for nonconvulsive status epilepticus. *Epilepsia* 2013;54(Suppl. 6):28–9.
- Hirsch LJ, LaRoche SM, Gaspard N, Gerard E, Svoronos A, Herman ST, et al. American Clinical Neurophysiology Society's standardized critical care EEG terminology: 2012 version. *J Clin Neurophysiol* 2013;30:1–27.