scientific reports

Check for updates

OPEN

The value of linear and non-linear quantitative EEG analysis in paediatric epilepsy surgery: a machine learning approach

Mattia Mercier[1,2], Chiara Pepi[1], Giusy Carfi-Pavia[1], Alessandro De Benedictis[3], Maria Camilla Rossi Espagnet[4], Greta Pirani[5], Federico Vigevano[1], Carlo Efisio Marras[3], Nicola Specchio[1✉] & Luca De Palma[1]

Epilepsy surgery is effective for patients with medication-resistant seizures, however 20–40% of them are not seizure free after surgery. Aim of this study is to evaluate the role of linear and non-linear EEG features to predict post-surgical outcome. We included 123 paediatric patients who underwent epilepsy surgery at Bambino Gesù Children Hospital (January 2009–April 2020). All patients had long term video-EEG monitoring. We analysed 1-min scalp interictal EEG (wakefulness and sleep) and extracted 13 linear and non-linear EEG features (power spectral density (PSD), Hjorth, approximate entropy, permutation entropy, Lyapunov and Hurst value). We used a logistic regression (LR) as feature selection process. To quantify the correlation between EEG features and surgical outcome we used an artificial neural network (ANN) model with 18 architectures. LR revealed a significant correlation between PSD of alpha band (sleep), Mobility index (sleep) and the Hurst value (sleep and awake) with outcome. The fifty-four ANN models gave a range of accuracy (46–65%) in predicting outcome. Within the fifty-four ANN models, we found a higher accuracy (64.8% ± 7.6%) in seizure outcome prediction, using features selected by LR. The combination of PSD of alpha band, mobility and the Hurst value positively correlate with good surgical outcome.

Resective surgery is highly effective in selected subgroups of patients with medication-resistant seizures[1,2], and between 40 and 80% of patients achieve complete post-operative seizure freedom (SF)[1–3]. Aetiology is the main predictor of post-surgical outcome[3]. A better understanding of the possible predictive value of pre-surgical investigations may improve surgical outcome[4,5]. A comprehensive diagnostic work-up with information coming from multiple modalities as electroencephalography (EEG), magnetic resonance imaging (MRI), neuroimaging post-processing and functional neuroimaging may estimate the likelihood of post-surgical seizure freedom, even if the weight of each investigation in determining the outcome is not completely understood[6,7]. Patients with clearly localized epileptogenic zone and concordance between multiple diagnostic modalities have a higher rate of seizure freedom[8]. Scalp EEG is a standard procedure used to localize the epileptogenic zone[2], suffering from discordant results[7,9]. Quantitative analysis of pre-surgical scalp EEG may help in identifying patients with higher likelihood of seizure freedom, even if such a methodologic approach is still underused[10] . The use of quantitative scalp EEG analysis is challenging due to the common presence of background noise, artefacts and the signal's non-stationarity nature[11].

[1]Neurology, Epilepsy and Movement Disorders Unit, Bambino Gesù Children's Hospital, IRCCS, Full Member of European Reference Network EpiCARE, Piazza S. Onofrio 4, 00165 Rome, Italy. [2]Department of Physiology, Behavioural Neuroscience PhD Program, Sapienza University, Rome, Italy. [3]Neurosurgery Unit, Bambino Gesù Children's Hospital, IRCCS, Full Member of European Reference Network EpiCARE, 00165 Rome, Italy. [4]Neuroradiology Unit, Imaging Department, Bambino Gesù Children's Hospital, 00165 Rome, Italy. [5]Department of Mechanical and Aerospace Engineering – DIMA, Sapienza University of Rome, Rome, Italy. ✉email: nicola.specchio@opbg.net

Traditional EEG analysis methods primarily focus on linear features such as amplitude and frequency, which have limited ability to capture the complex dynamics of the signal[12]. Non-linear EEG features, like Lyapunov Exponent and Entropy, provide a more comprehensive analysis of the EEG signal by capturing the underlying dynamics of the signal[13]. Interictal linear and non-linear EEG analysis may disclose pro-epileptogenic regions[14–16]. These analysis are mainly used to predict epileptic seizures and not surgical outcome[17]. Entropy[18,19] and Lyapunov exponents[20], can be used to quantify the level of epileptogenicity examining the brain's predictability and chaos levels.

Non-linear EEG features derived from EEG rhythms and wavelet decomposition, achieve a high prediction accuracy (85–93%) in detecting focal versus non-focal epilepsies[21]. Other linear and non-linear biomarkers, such as Hjorth parameters[22] and Hurst index[23,24], may provide an estimate of the dynamic interaction in different scalp EEG signals, and may predict seizure lateralization[22]. Power spectral density, hjorth, approximate entropy, permutation entropy, Lyapunov and Hurst exponent in scalp EEG are limited by artefacts, low spatial resolution and low sampling[25].

Brain machine learning (BML) is a powerful tool in biomedical applications often underused in EEG analysis[26]. BML methods are used for automated epilepsy diagnosis prediction[27], seizure detection (accuracy 57–100%)[28] and localization (accuracy 93–96%)[29].

No studies have demonstrated how BML algorithms can be employed to evaluate the role of a specific combination of linear and non-linear interictal EEG features to predict surgical outcome.

Aim of this study is uncover hidden EEG features that may contribute to the variability of epilepsy surgery outcomes, computing and integrating linear and non-linear features in awake and sleep scalp and interictal EEG signals and applying a BML approach.

Results

We screened 246 interictal scalp recordings (2 EEGs for each patient, one recording during wakefulness and one during sleep) from 123 patients. All EEGs were free from artifact and epileptiform abnormalities. We collected 1-min of each EEG to extract linear and nonlinear features. The descriptive analysis of the linear and nonlinear EEG features is summarized in Supplementary Table 1. The descriptive comparison of non-linear EEG features are represented in Fig. 1.

The results of logistic regression (LR) test showed that Hurst value during sleep EEG is positively correlated with seizure freedom (OR = 2.681, 1.084 < CI < 6.629, $p = 0.033$), while Hurst value during wakefulness EEG was negatively correlated with seizure freedom (OR = 0.281, 0.082 < CI < 0.967, $p = 0.044$).

Power spectral density (PSD) in alpha band during sleep (OR = 1.400, 1.001 < CI < 1.007, $p = 0.019$) and mobility index during sleep (OR = 2.783, 1.140 < CI < 6.797, $p = 0.025$) were positively correlated with seizure freedom. These statistically significant EEG features were used in SET 3. No statistical correlation was found between Hjorth, approximate entropy, permutation entropy, and Lyapunov indexes and surgical outcome ($p > 0.05$).

The results of accuracy of each artificial neural network (ANN) using SET 1 (all features), SET 2 (averaged features) and SET 3 (features selected by LR) as training sets are reported in Table 1. The comparison of prediction accuracy (P%) generated by different ANN architectures (I-IX) and SET (1–3) of features is shown in Fig. 2.
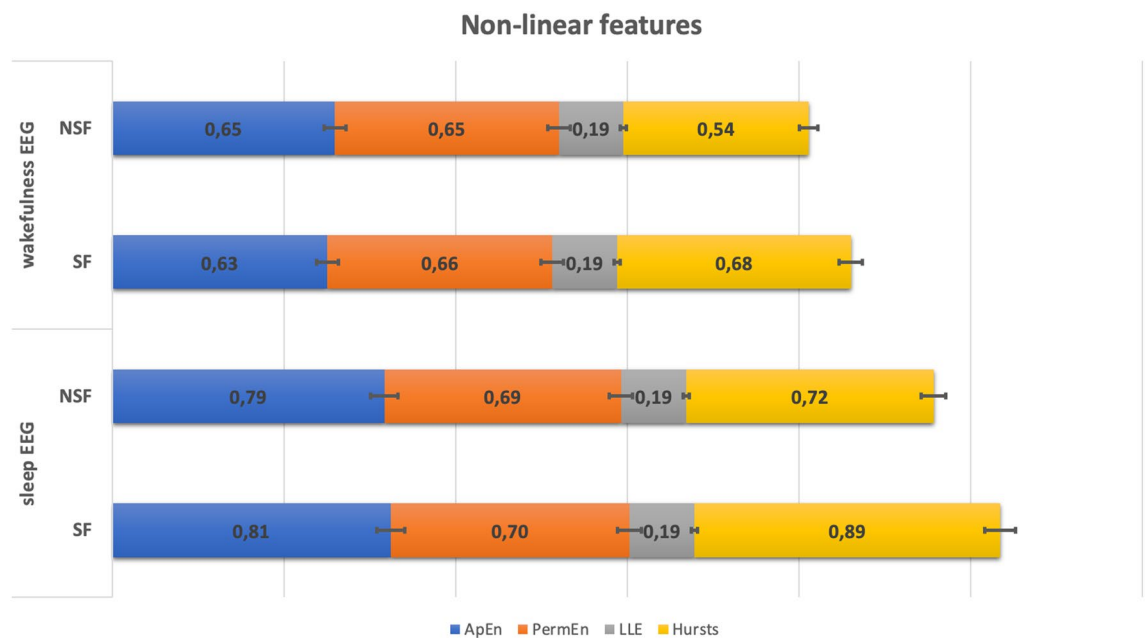
Figure 1. Mean values of Hurst, ApEn, PermEn and LLE non-linear EEG features extracted for both sleep and wakefulness condition: comparison between SF and NSF surgery outcome values. ApEn, approximate entropy; PermEn, permutation entropy; LLE, Lyapunove Exponent.

Sets	Architecture	Topology A				Topology B			
		Nmax	Accuracy P (%)	Sensibility (%)	Specificity (%)	Nmax	Mean accuracy P (%)	Sensibility (%)	Specificity (%)
Set1	I	45	55.2 ± 7.7	87.3 ± 7.3	23.0 ± 14.2	40	52.7 ± 9.3	79.3 ± 13.5	26.0 ± 15.1
	II	45 + 45	52.2 ± 6.6	89.3 ± 10.5	15.0 ± 1.1	40 + 20	48.0 ± 6.9	80.0 ± 14.7	16. 0 ± 9.7
	III	45 + 45 + 45	50.8 ± 4.5	90.7 ± 10.5	11.0 ± 9.9	40 + 20 + 10	52.0 ± 5.5	88.0 ± 13.6	16.0 ± 19.6
	IV	40	50.2 ± 8.7	85.3 ± 5.3	15.0 ± 16.5	30	51.3 ± 7.7	90.7 ± 15.1	12.0 ± 19.9
	V	40 + 40	46.3 ± 6.5	82.7 ± 11.8	10.0 ± 6.7	30 + 15	47.2 ± 7.7	81.3 ± 15.7	13.0 ± 6.7
	VI	40 + 40 + 40	48.0 ± 3.3	84.0 ± 9	12.0 ± 13.2	30 + 15 + 8	53.7 ± 7.2	81.3 ± 10.8	26.0 ± 17.8
	VII	35	52.7 ± 7.8	77.3 ± 14.1	28.0 ± 11.4	25	49.5 ± 4.4	80.0 ± 16.3	19.0 ± 11
	VIII	35 + 35	51.8 ± 10.4	80.7 ± 13.5	23.0 ± 18.9	25 + 12	54.0 ± 7.3	82.0 ± 17.5	26.0 ± 20.1
	IX	35 + 35 + 35	53.7 ± 3.3	83.3 ± 16.4	24.0 ± 20.7	25 + 12 + 6	51.5 ± 5.2	86.0 ± 10.2	17.0 ± 8.2
Set2	I	36	53.5 ± 8.7	84.0 ± 11.4	23.0 ± 18.3	36	51.0 ± 9.7	78.0 ± 14.4	24.0 ± 11.7
	II	36 + 36	55.0 ± 13.3	82.0 ± 13.7	28.0 ± 22	36 + 18	49.8 ± 5.7	78.7 ± 15.7	21.0 ± 16
	III	36 + 36 + 36	53.5 ± 6.2	80.0 ± 16.6	27.0 ± 17	36 + 18 + 9	49.2 ± 7.4	87.3 ± 15.5	11.0 ± 7.4
	IV	30	52.0 ± 5.7	82.0 ± 13.7	22.0 ± 13.2	30	47.8 ± 6.6	70.7 ± 21.4	25.0 ± 23
	V	30 + 30	53.5 ± 8.1	80.0 ± 18.1	27.0 ± 15.7	30 + 15	52.7 ± 7.7	87.3 ± 11.5	18.0 ± 14.8
	VI	30 + 30 + 30	53.3 ± 11.9	76.7 ± 11.9	30.0 ± 20	30 + 15 + 8	52.5 ± 11	80.0 ± 18.3	25.0 ± 17.2
	VII	20	52.2 ± 10	83.3 ± 15.8	21.0 ± 12	20	53.8 ± 6.2	80.7 ± 18.4	27.0 ± 21.1
	VIII	20 + 20	51.2 ± 10.2	85.3 ± 14	17.0 ± 15.7	20 + 10	49.0 ± 6	80.0 ± 12.6	18.0 ± 12.3
	IX	20 + 20 + 20	52.5 ± 9.7	88.0 ± 12.1	17.0 ± 20	20 + 10 + 5	50.3 ± 12.2	70.7 ± 10.5	30.0 ± 19.4
Set3	I	9	59.2 ± 7.9	79.3 ± 15.2	25.0 ± 15.8	9	58.8 ± 7.1	76.7 ± 14.8	31.0 ± 18.5
	II	9 + 9	64.8 ± 7.6	76.7 ± 14.1	43.0 ± 14.2	9 + 5	59.3 ± 7	84.7 ± 15.1	20.0 ± 18.9
	III	9 + 9 + 9	62.0 ± 7.9	76.0 ± 11.8	24.0 ± 19.6	9 + 5 + 3	59.2 ± 7.7	61.3 ± 17.7	43.0 ± 26.3
	IV	6	61.0 ± 6.4	80.0 ± 11.8	22.0 ± 18.7	6	55.3 ± 9.8	84.7 ± 13.4	26.0 ± 16.5
	V	6 + 6	59.2 ± 6.8	79.3 ± 15.5	21.0 ± 12.9	6 + 3	53.3 ± 5.6	80.7 ± 27.7	26.0 ± 32
	VI	6 + 6 + 6	59.1 ± 6.2	67.3 ± 11.1	31.0 ± 11	6 + 3 + 2	52.2 ± 6.4	91.3 ± 8.3	13.0 ± 17
	VII	4	61.8 ± 6.8	84.7 ± 7.7	19.0 ± 9.9	4	55.2 ± 6.2	87.3 ± 7.3	23.0 ± 11.6
	VIII	4 + 4	57.0 ± 9.9	82.0 ± 16.6	32.0 ± 18.1	4 + 2	54.5 ± 10.4	82.0 ± 16	27.0 ± 27.5
	IX	4 + 4 + 4	58.0 ± 10.1	82.0 ± 9.5	24.0 ± 18.54	4 + 2 + 1	53.0 ± 7.7	78.0 ± 17.2	28.0 ± 21

Table 1. Results of accuracy, sensibility and specificity of the nine artificial neural networks architectures, topology and SET of features used as input of models. The best classification performance is related to Set 3-II ANN architecture, topology A (64.8%). The best model had a 76.7% sensitivity and 43% of specificity, which was higher than other architectures tested. Nmax, maximum number of neurons in each hidden layer used.

Analysing the results of BML (Table 1 and Fig. 2), we found that the II architecture of ANN was the most accurate in predicting seizure outcome, with a specific topology of 2HL and 9 N, and a specific set of EEG features (SET 3). This combination resulted in the highest prediction accuracy ($p = 64.8\% \pm 7.6$). The optimal combination of EEG features is presented as linear combination of 4 linear and non-linear EEG features, including Hurst during wakefulness and sleep recordings, Mobility during sleep EEG, and PSD in alpha band during sleep EEG.

SET 3, represented by the light-blue bar, emerged as the most effective training set across all tested ANN architectures, as showed in Fig. 2. When combined with topology A (Fig. 1), SET 3 achieved superior predictive accuracy for all ANNs architectures, with accuracy values ranging from 57 to 64.8%. The use of SET 3 outperformed both SET 1 and SET 2. The analysis revealed no notable differences in predictive performance between SET 1 and SET 2. Similarly, no significant distinctions were observed among the various topologies of the ANNs. ($p > 0.05$, CI = 95%).

The best model, featuring II ANN architecture with topology A and SET 3, demonstrated superior prognostic accuracy in identifying SF outcomes, achieving a notable sensitivity of 76.7% and a specificity of 43%, exceeding the performance metrics of all other models tested.

The specificity of best model was the highest compared to all architectures using SET 3.

Comparing the performance of different architectures (from I to IX) related to the best model (II architecture, SET 3 and topology A) we found that the accuracy was significantly higher than all other combinations of topology and SET (p value < 0.05, CI = 95%).

Discussion

We explored the role of EEG dynamic properties in predicting epilepsy surgery outcome. Currently only clinical, neuroimaging and neurophysiological qualitative predictors[1,3,9] are strongly correlated with post-surgical seizure outcome, moreover only some of these predictors are used to generate predictive models of seizure outcome[9]. To characterize EEG segments, we composed three different sets (SET1, SET2, SET3) of EEG features as descripted in Methods section. We applied two different approaches. To compute SET 1 we performed a single channel extraction: all predicting values are computed at each channel of standard 10–20 montage and epoch as reported

3
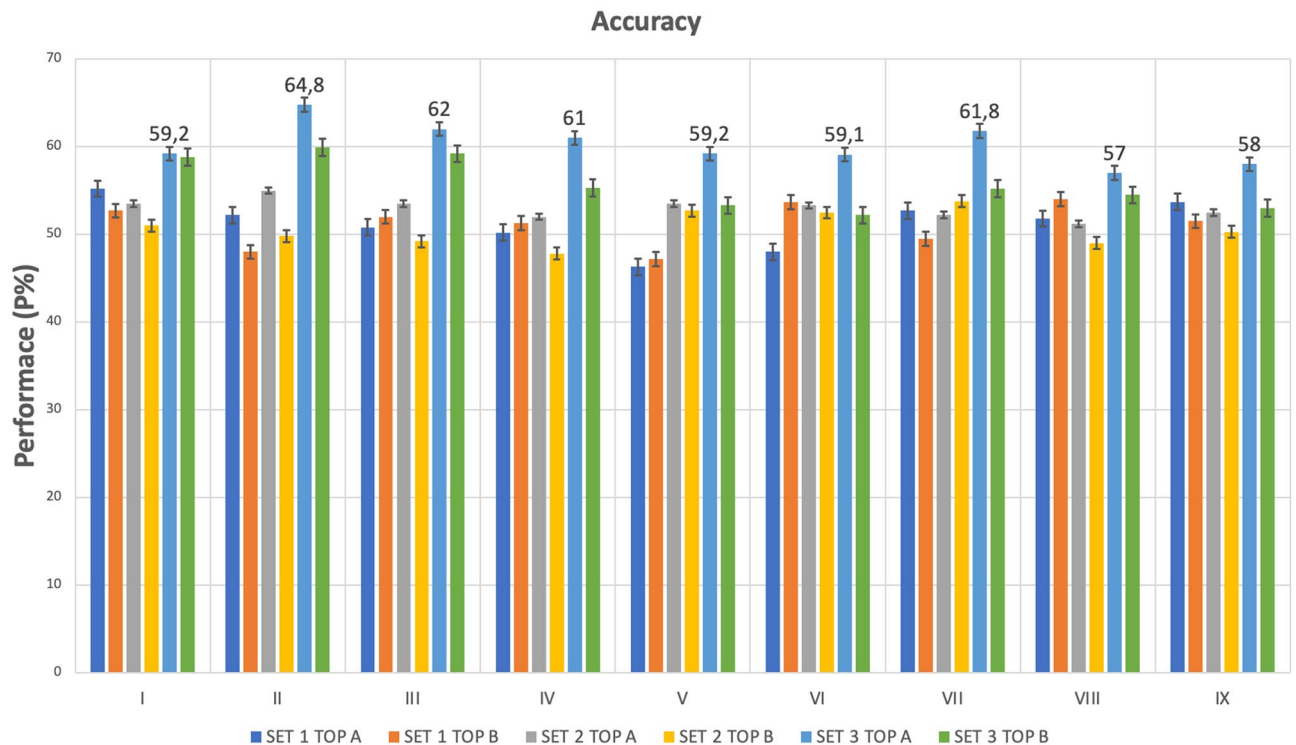
Figure 2. Comparison of prediction accuracy (P%) generated by different ANN architectures (I–IX), SET of features (1–3) and ANN topology (A–B). The best accuracy results are related to Set 3-II ANN architecture, Topology A and SET 3 (light-blue bar) were the best combination of SET and topology in terms of accuracy. SET 3 with topology A led to a higher predictive value for all ANNs (P% = 57–64.8%) compared to SET 1 and SET 2.

in Lemoine et al.[30]. To compute SET 2 and 3 we extracted all values as the average across all channels according to Lin et al.[31,32]. This approach may mitigate the variability in single-channel data.

We used selected features for the SET 3 which were extracted using LR, and we achieved the best accuracy in predicting surgical outcome. Among power spectral density, Hjorth, approximate entropy, permutation entropy, Lyapunov and Hurst exponent averaged EEG features, some may be redundant or may not contain enough discriminative information for the prediction[33]. In addition, the ANN with a high number of EEG features may be influenced by the relatively low number of patients[34].

Hurst exponent is the only non-linear EEG feature able to discriminate between SF and non-Seizure freedom (NSF) patients. LR analysis demonstrated that increasing the regularity (increasing Hurst value) of the EEG signal during sleep, the chances of seizure freedom increased (OR = 2.681, 1.084 < CI < 6.629 with p = 0.033), while increasing the regularity of the EEG signal during wakefulness, the chances of seizure freedom decreased (OR = 0.281, 0.082 < CI < 0.967 with p = 0.044).

The LR considers the distribution of values and their relationship with the outcome variable, rather than a simple comparison of mean values. An increase in Hurst values in wakefulness is associated with a lower probability of attaining seizure freedom. On the other hand in sleep a "positive correlation" is an increase in Hurst values associated with a higher probability of achieving seizure freedom. A study by Witton et al.[35] analysed Hurst exponent of pre-surgical EEG signals and found that Hurst value was able to identify the probable epileptogenic zone in 3 out of 3 patients (100%). The interpretation of Hurst values can be challenging, as they are affected by signal length, noise level, and sampling rate[36].

Alpha band PSD is a potential biomarker for the automatized detection of epileptic seizures, achieving a 98% of accuracy model[37], even if alpha band PSD analysis of EEG signals is affected by the total power of the spectrum[38]. In our study Alpha band PSD during sleep stage is positively correlated with seizure outcome (OR = 1.400, 1.001 < CI < 1.007 with p = 0.019). No previous studies demonstrated the value of alpha band PSD in predicting epilepsy surgical outcome in a paediatric population, differently changes in the alpha band PSD have been correlated with several neurological and psychiatric diseases in adults[39].

Mobility had only been studied in seizure prediction and lateralization[40]. In addition, C.S Ouyang et al.[41] observed a significant increase of mobility in patients who benefit from anti-seizure medications. In the present study, the mobility index calculated during the sleep state was correlated with the post-surgical outcome indicating that higher value significantly improves the probability of seizure freedom (OR = 2.783, 1.140 < CI < 1.6.797 with p = 0.025). These results confirm our previous study showing that the mobility index positively correlated with favourable surgical outcome in patients undergoing hemispherectomy (p = 73%)[42].

We then focused on the specific combination of linear and non-linear EEG features to predict surgical outcome.

Lemoine et al. investigated how combining linear and non-linear EEG features could predict seizure recurrence within 1 year after EEG, using four BML algorithms (general linear model, support vector machine, Random Forest and LightGBM). They achieved an accuracy rate between 62 and 67%[30]. Previously, no studies had showed that such a combination of linear and non-linear interictal scalp EEG features could accurately predict surgical outcome in children with epilepsy. In our three SETs of EEG features, SET 3 showed the most promising results revealing the lowest mean square error (MSE = 9.2). This indicates that our choice of features and the size of our dataset were well-suited for creating a stable model[30,42,43].

In the last few years more and more studies tried to predict the post-surgical outcome using predictive models. The best results were achieved using clinical variables. Grigsby et al. trained an ANN classifier using clinical, neuropsychological and imaging data from 65 patients treated with anterior temporal lobectomy; the accuracy was of 81.8% in predicting Engel I outcome (improving to 95.4% for Engel I or II outcomes)[4]. Arle et al. also applied ANN with several architecture, reporting an accuracy of 96% in predicting Engel I outcomes in unselected 80 surgical patients[44]. Different methods have been utilized (nomograms and simple seizure-freedom scores) to predict seizure freedom in mixed adult and paediatric populations[1,45] with poor predictive value (AUC of 0.528–0.539 and 0.533–0.539 at 2 and 5 years time points)[46]. Sinclair et al. evaluated, also, the potential of BML techniques applied to standard presurgical brain MRs and PET scans to provide enhanced prognostic value to such neuroimaging tools. Up to 73% of patients with poor surgical outcome were predicted, potentially providing additional information to incorporate into surgical decision-making process[47].

The choice of specific BML tool and the number of architectures and topology are still not properly defined. We do not have a pre-defined model with fixed number of architectures. Our best model (II—SET3—topology A) showed a significant higher accuracy than all other combinations of topology and SETs (p value < 0.05). We observed that, using SET 3, the ANN performed better than using SET 1 (all features) and SET 2 (averaged), as it is shown in Table 1. The choice to use 3 training SETs was arbitrary. Each set of features might involve different levels of processing leveraging the strengths of each of them[48].

Our findings suggest that a specific architecture and specific selection of EEG features may improve ANN performance, indicating that not all EEG features are effective in predicting epilepsy surgical outcome. Similar results were found in a previous study which demonstrated that some clinical and EEG features are irrelevant for prognosis. In this study three BML models (Naïve Bayes, logistic regression and K-NN) were used: authors extracted from 23 patients with Hippocampal Sclerosis a specific set of features achieving an improvement of accuracy from 68.42 to 89.47%[49].

Accurate prediction of seizure outcome after epilepsy surgery remains difficult. traditional statistical modelling (LR) and machine learning techniques (multilayer perceptron and XGBoost) performed equally (72% vs 71%, $p > 0.05$) to predict 1-year post-operative seizure outcome on 797 children who undergone resective or disconnective surgery[50].

We do have some limitations to be acknowledged. We were missing an external validation cohort, and this could lead to underperform if the same model is tested on data coming from a different sample[50]. Despite this limitation, the ANN may provide a powerful tool to optimize patient management[51,52] improving the inherent characteristics and quality of data. As previously recommended[50], we also strongly believe that a collaboration to create standardized datasets, selection of appropriate predictor variables for modelling, sharing of models and code, are essential for advancing this research field. It is important to note that in our study we used only scalp EEG signals, which are known to have lower spatial resolution and high level of signal noise[53,54], if compared with to Stereo-EEG signals. Moreover, the performance of the ANN model may be affected by the dataset size and the monocentric recruitment. Future studies may also consider interpolating other types of data, such as neuroradiological and clinical features, to improve the accuracy of the ANN. Furthermore, our study included only interictal scalp EEG segments free from epileptiform abnormalities, and it may turn that EEG signal with epileptiform abnormalities can discriminate better dynamical EEG properties[55]. Despite these limitations, our results suggest that this ANN model may hold considerable promise as adjuncts to clinical expertise and not as a replacement.

Our study is the first to investigate the relationship between linear and non-linear EEG properties and surgical outcomes in paediatric epilepsy patients. We found that a specific combination of EEG features, such as the Hurst exponent, Mobility Index and PSD, were correlated with post-surgical seizure freedom, achieving an accuracy of 64.8% in predicting surgical outcome. The main contributions of this study are: (1) the first development of an automated and quantitative approach and tool for early prediction of epilepsy based on interictal EEG classification analysis; (2) identification of significant linear and non-linear EEG features for discriminating between SF and NSF patients.

Methods

We retrospectively reviewed all paediatric patients who underwent surgery for medication-resistant epilepsy from January 2009 to April 2020, at Bambino Gesù Children's Hospital, Rome, Italy. All methods were carried out in accordance with relevant guidelines and regulations. Ethical approval and inform consent were waiver by ethical committee of the Ospedale Bambino Gesu'. Data were retrospectively analysed in line with personal data protection policies.

We found 169 patients. We included only patients with the following inclusion criteria:

(1) Available pre-surgical EEG data during wakefulness and sleep.
(2) One-minute artefact free awake and sleep EEG.
(3) Post-surgical follow-up of at least 3 years.

One-hundred-twenty-three patients were enrolled in the study. The mean age at surgery is 7.3 ± 12.2 years, 73 out of 123 (59.4%) patients are seizure free and drug free (SF) and never experienced post-surgical seizures; 50 out of 123 (40.6%) are non-seizure free (NSF). Supplementary Table 2 show the results of histopathology examination on brain specimen. All patients underwent routine pre-surgical evaluation, including full history and neurological examination, brain MRI and visual analysis of long-term video-EEG monitoring. All patients underwent neuropsychological assessment during follow-up.

The study includes six stages (Fig. 3): EEG recording, signal processing and analysis, features extraction and selection, and classification. EEG recordings were obtained with a VEEG monitoring system (Micromed, Treviso, Italy) at the Neurology, Epilepsy and Movement Disorders Unit of the Bambino Gesù Children's Hospital in Rome, Italy. The signal processing, analysis and classification were computed with MATLAB software (R2022b). The 10–20 electrode montage was used for scalp recordings.

The reference electrode was set as the average of all contacts. The monopolar recordings were obtained with a sampling frequency of 256 and 512 Hz, powerline notch filtered at 50 Hz, band-pass filtered between 0.5 and 45 Hz (7th order Butterworth filter) and 16-bit resolution and z-score standardized.

The extraction of EEG data was performed primarily by neurophysiology expert (CP, GCP) through visual inspection. Before filtering, the EEG signals were down sampled (256 Hz). Sixty seconds of EEG signal was used both in wakefulness and sleep. The EEG data were without artefact and spikes[42]. EEG feature extraction was performed based on a sliding-window approach. The size of the window (l) was long enough to capture temporal patterns of the signal[56], while considering the assumption of stationarity of the time series. The size of the window (l) is set pair to 5 or 10 s considering the different EEG features but is never below 60 s. We extracted 13 linear and non-linear features for each EEG signal using non-overlapping windows approach. We collected the data from all the 19 EEG electrodes for each patient during both wakefulness and sleep.

Methodology in brain machine learning approach

The EEG features are used as input to the ANN classifier. We used a linear and non-linear methods to analyse data[25] as reported in Supplementary Methods. An ANN approach was used for prediction of outcome after surgery using linear and non-linear features extracted from the pre-surgery EEG data as is shown in the Fig. 3. We trained 2 different topologies of ANN (A–B) and 9 different architectures (I–IX) of feedforward network with different numbers of hidden layers (HL) and different number of neurons (N) in each HL. The number of HL varied in the range of 1–3, while the number N in each HL varied based on the number of N in the first hidden layer (Tab. 2). We set a total of 54 confusion matrices, eighteen for each input SET. The maximum number of N was set following the empirical formula developed by Yotov et al.[57]. The number of neurons vary depending on the size of the training set. The output set consisted of two coded values, SF = 1 and NSF = 0. All networks were trained with a supervised approach using the Conjugate Gradient[58]. To verify the reproducibility of our results, all networks were trained 20 times using 70% of input patients randomly chosen as the training test, a random 15% of patients as the validation set and a random 15% of patients as testing set. A cross-validation scheme was used to train and test each classifier: the prediction accuracy was computed as the average of twenty iterations. To prevent overfitting and to improve performance and generalizability in the BML model we performed a feature selection: we divided our features in 3 different training sets, defined as a linear combination of different EEG features. We used the following features dataset:

(1) SET 1: all features acquired by each channel;
(2) SET 2: the average of all channel's EEG features;
(3) SET 3: EEG features statistically significant to the LR test after evaluating the correlation between the EEG pre-surgical features with the epileptic outcome (SF vs NSF).
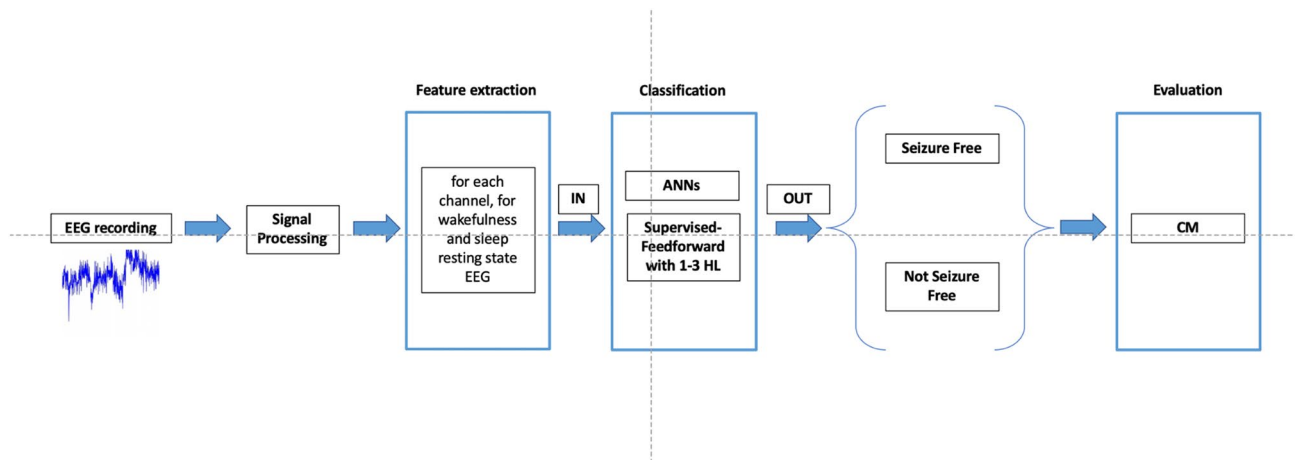
Figure 3. Block-diagram of the proposed surgical outcome prediction method using artificial neural network (ANN). HL, hidden layers; CM, confusion matrix.

Architectures	Topology A	Topology B
I	n1	n1
II	n1 + n1	n1 + 1/2n1
III	n1 + n1 + n1	n1 + 1/2n1 + 1/3n1
IV	n2	n2
V	n2 + n2	n2 + 1/2n2
VI	n2 + n2 + n2	n2 + 1/2n2 + 1/3n2
VII	n3	n3
VIII	n3 + n3	n3 + 1/2n3
IX	n3 + n3 + n3	n3 + 1/2n3 + 1/3n3

Table 2. Different topology of ANN based on nine different artificial neural network architectures (from I to IX). They were defined by a specific combination of hidden layers (HL) and specific number of neurons (N).

Odds ratio (OR) in LR model was used to study the positive or negative correlation between the pre-surgical EEG features and the post-surgical outcome (SF *vs* NSF) to define the EEG features of SET 3; *p*-value below 0.05 were considered statistically significant. In implementing the LR model, the selected "response event" is always the SF condition. The set of features and the maximum number of neurons for each test is reported in Supplementary Table 3.

For each trained network, a confusion matrix was calculated based on the real output (seizure free or non-seizure free) and the one estimated on the randomly extracted testing set.

The mean 2×2 confusion matrix was then obtained by averaging the confusion matrixes of the trained ANNs for each iteration. A performance parameter (P) was calculated as the mean (%) of the elements on the diagonal of the mean confusion matrix, where 100% indicates the absence of misclassifications (Table 2).

The mean square error (MSE) was calculated to select the most accurate training sets. The whole analysis process is illustrated in the Fig. 4.

The Wilcoxon Signed-Rank Test was used to rank the differences in performance between the best accurate classifier and the other ones for each architecture, and the *p*-values were found to be less than 0.05 with a confidence interval of 95%. We did not apply a *p*-value correction in the analyses stems from the exploratory nature of the study.
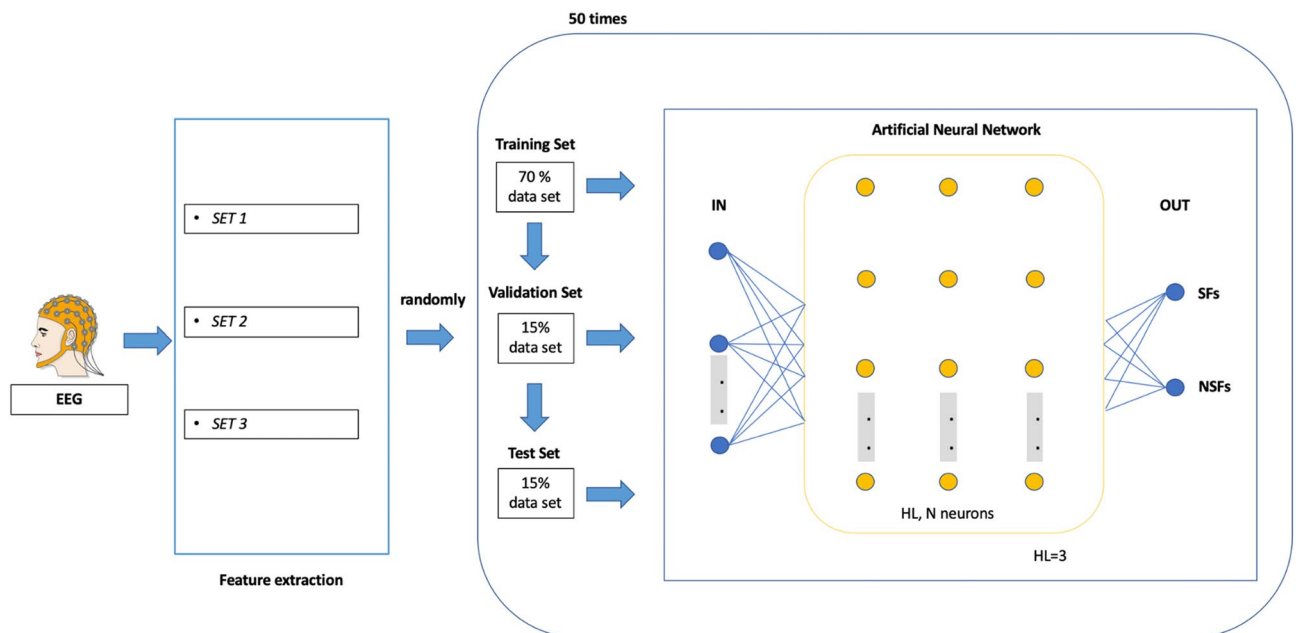
Figure 4. Study workflow and ANN architectures of the model. We used 3 Sets of EEG features to train and test the ANN models. We selected the 70% of data as training set and 30% for the validation and test sets. All architectures are composed of specific N and HL related to input data. Surgical outcome is dichotomous (SF–NSF). HD, hidden layer; N, neurons; SF, seizure freedom and drug freedom, NSF, not seizure freedom.

Data availability

We declare that data supporting the results of this study are stored at Bambino Gesu' Children's Hospital internal repository and are available upon request to the corresponding author Dr. Nicola Specchio.

References

1. Jehi, L. *et al.* Development and validation of nomograms to provide individualised predictions of seizure outcomes after epilepsy surgery: A retrospective analysis. *Lancet Neurol.* **14**, 283–290. https://doi.org/10.1016/S1474-4422(14)70325-4 (2015).
2. Cross, J. H., Reilly, C., Delicado, E. G., Smith, M. L. & Malmgren, K. Epilepsy surgery for children and adolescents: Evidence-based but underused. *Lancet Child Adolesc. HealTH* **6**, 484–494. https://doi.org/10.1016/S2352-4642(22)00098-0 (2022).
3. Lamberink, H. J. *et al.* Seizure outcome and use of antiepileptic drugs after epilepsy surgery according to histopathological diagnosis: A retrospective multicentre cohort study. *Lancet Neurol.* **19**, 748–757. https://doi.org/10.1016/S1474-4422(20)30220-9 (2020).
4. Grigsby, J., Kramer, R. E., Schneiders, J. L., Gates, J. R. & Brewster, S. W. Predicting outcome of anterior temporal lobectomy using simulated neural networks. *Epilepsia* **39**, 61–66. https://doi.org/10.1111/j.1528-1157.1998.tb01275.x (1998).
5. Jehi, L. Machine learning for precision epilepsy surgery. *Epilepsy Curr.* **23**, 78–83. https://doi.org/10.1177/15357597221150055 (2023).
6. Memarian, N., Kim, S., Dewar, S., Engel, J. J. & Staba, R. J. Multimodal data and machine learning for surgery outcome prediction in complicated cases of mesial temporal lobe epilepsy. *Comput. Biol. Med.* **64**, 67–78. https://doi.org/10.1016/j.compbiomed.2015.06.008 (2015).
7. Alim-Marvasti, A., Vakharia, V. N. & Duncan, J. S. Multimodal prognostic features of seizure freedom in epilepsy surgery. *J. Neurol. Neurosurg. Psychiatry* **93**, 499–508. https://doi.org/10.1136/jnnp-2021-327119 (2022).
8. Najm, I. *et al.* Temporal patterns and mechanisms of epilepsy surgery failure. *Epilepsia* **54**, 772–782. https://doi.org/10.1111/epi.12152 (2013).
9. Fitzgerald, Z. *et al.* Improving the prediction of epilepsy surgery outcomes using basic scalp EEG findings. *Epilepsia* **62**, 2439–2450. https://doi.org/10.1111/epi.17024 (2021).
10. Varatharajah, Y. *et al.* Quantitative analysis of visually reviewed normal scalp EEG predicts seizure freedom following anterior temporal lobectomy. *Epilepsia* **63**, 1630–1642. https://doi.org/10.1111/epi.17257 (2022).
11. Huang, J. S., Li, Y., Chen, B. Q., Lin, C. & Yao, B. An intelligent EEG classification methodology based on sparse representation enhanced deep learning networks. *Front. Neurosci.* **14**, 1–8. https://doi.org/10.3389/fnins.2020.00808 (2020).
12. Morales, S. & Bowers, M. E. Time-frequency analysis methods and their application in developmental EEG data. *Dev. Cogn. Neurosci.* **54**, 101067. https://doi.org/10.1016/j.dcn.2022.101067 (2022).
13. Stam, C. J. Nonlinear dynamical analysis of EEG and MEG: Review of an emerging field. *Clin. Neurophysiol.* **116**, 2266–2301. https://doi.org/10.1016/j.clinph.2005.06.011 (2005).
14. Müller, M. *et al.* Linear and nonlinear interrelations show fundamentally distinct network structure in preictal intracranial EEG of epilepsy patients. *Hum. Brain Mapp.* **41**, 467–483. https://doi.org/10.1002/hbm.24816 (2020).
15. Acharya, U. R., Fujita, H., Sudarshan, V. K., Bhat, S. & Koh, J. E. W. Application of entropies for automated diagnosis of epilepsy using EEG signals: A review. *Knowl. Based Syst.* **88**, 85–96. https://doi.org/10.1016/j.knosys.2015.08.004 (2015).
16. Lehnertz, K. Epilepsy and nonlinear dynamics. *J. Biol. Phys.* **34**, 253–266. https://doi.org/10.1007/s10867-008-9090-3 (2008).
17. Andrzejak, R. G. *et al.* Improved spatial characterization of the epileptic brain by focusing on nonlinearity. *Epilepsy Res.* **69**, 30–44. https://doi.org/10.1016/j.eplepsyres.2005.12.004 (2006).
18. Travnicek, V. *et al.* Relative entropy is an easy-to-use invasive electroencephalographic biomarker of the epileptogenic zone. *Epilepsia* **64**(4), 962–972. https://doi.org/10.1111/epi.17539 (2023).
19. Staniek, M. & Lehnertz, K. Parameter selection for permutation entropy measurements. *Int. J. Bifurc. Chaos* **17**, 3729–3733. https://doi.org/10.1142/S0218127407019652 (2007).
20. Sathyanarayana, A. *et al.* Nonlinear analysis of visually normal EEGs to differentiate benign childhood epilepsy with centrotemporal spikes (BECTS). *Sci. Rep.* **10**, 1–12. https://doi.org/10.1038/s41598-020-65112-y (2020).
21. Akbari, H. & Sadiq, M. T. Detection of focal and non-focal EEG signals using non-linear features derived from empirical wavelet transform rhythms. *Phys. Eng. Sci. Med.* **44**, 157–171. https://doi.org/10.1007/s13246-020-00963-3 (2021).
22. Cecchin, T. *et al.* Seizure lateralization in scalp EEG using Hjorth parameters. *Clin. Neurophysiol.* **121**, 290–300. https://doi.org/10.1016/j.clinph.2009.10.033 (2010).
23. Lahmiri, S. Generalized Hurst exponent estimates differentiate EEG signals of healthy and epileptic patients. *Phys. A Stat. Mech. Its Appl.* **490**, 378–385. https://doi.org/10.1016/j.physa.2017.08.084 (2018).
24. Indiradevi, K. P., Elias, E. & Sathidevi, P. S. Complexity analysis of electroencephalogram records of epileptic patients using Hurst exponent. *Int. J. Med. Eng. Inform.* **1**, 368–380. https://doi.org/10.1504/IJMEI.2009.022647 (2009).
25. Subha, D. P., Joseph, P. K., Acharya, U. R. & Lim, C. M. EEG signal analysis: A survey. *J. Med. Syst.* **34**, 195–212. https://doi.org/10.1007/s10916-008-9231-z (2010).
26. Rajkomar, A., Dean, J. & Kohane, I. Machine learning in medicine. *N. Engl. J. Med.* **380**, 1347–1358. https://doi.org/10.1056/nejmra1814259 (2019).
27. Abbasi, B. & Goldenholz, D. M. Machine learning applications in epilepsy. *Epilepsia* **60**, 2037–2047. https://doi.org/10.1111/epi.16333 (2019).
28. Mirchi, N. *et al.* Decoding intracranial EEG with machine learning: A systematic review. *Front. Hum. Neurosci.* https://doi.org/10.3389/fnhum.2022.913777 (2022).
29. Daoud, H. & Bayoumi, M. Deep learning approach for epileptic focus localization. *IEEE Trans. Biomed. Circuits Syst.* https://doi.org/10.1109/TBCAS.2019.2957087 (2019).
30. Lemoine, É. *et al.* Machine-learning for the prediction of one-year seizure recurrence based on routine electroencephalography. *Sci. Rep.* **13**, 12650. https://doi.org/10.1038/s41598-023-39799-8 (2023).
31. Lin, L. C. *et al.* Early prediction of medication refractoriness in children with idiopathic epilepsy based on scalp EEG analysis. *Int. J. Neural Syst.* https://doi.org/10.1142/S0129065714500233 (2014).
32. Ouyang, C. S. *et al.* Quantitative EEG findings and response to treatment with antiepileptic medications in children with epilepsy. *Brain Dev.* **40**, 26–35. https://doi.org/10.1016/j.braindev.2017.07.004 (2018).
33. Qazi, E. H. *et al.* Single trial EEG patterns for the prediction of individual differences in fluid intelligence. *Front. Hum. Neurosci.* **10**, 1–19. https://doi.org/10.3389/fnhum.2016.00687 (2017).
34. Jia, W., Sun, M., Lian, J. & Hou, S. Feature dimensionality reduction: A review. *Complex Intell. Syst.* **8**, 2663–2693. https://doi.org/10.1007/s40747-021-00637-x (2022).
35. Witton, C. *et al.* Rogue bioelectrical waves in the brain: The Hurst exponent as a potential measure for presurgical mapping in epilepsy. *J. Neural Eng.* **16**, 056019. https://doi.org/10.1088/1741-2552/ab225e (2019).

36. Supriya, S., Siuly, S., Wang, H. & Zhang, Y. Automated epilepsy detection techniques from electroencephalogram signals: A review study. *Health Inf. Sci. Syst.* **8**, 1–15. https://doi.org/10.1007/s13755-020-00129-1 (2020).
37. Sameer, M. & Gupta, B. Detection of epileptical seizures based on alpha band statistical features. *Wirel. Pers. Commun.* **115**, 909–925. https://doi.org/10.1007/s11277-020-07542-5 (2020).
38. Park, T., Lee, M., Jeong, T., Shin, Y. I. & Park, S. M. Quantitative analysis of EEG power spectrum and EMG median power frequency changes after continuous passive motion mirror therapy system. *Sensors (Switzerland)* https://doi.org/10.3390/s20082354 (2020).
39. Larsson, P. G. & Kostov, H. Lower frequency variability in the alpha activity in EEG among patients with epilepsy. *Clin. Neurophysiol.* **116**, 2701–2706. https://doi.org/10.1016/j.clinph.2005.07.019 (2005).
40. Kang, J. H., Chung, Y. G. & Kim, S. P. An efficient detection of epileptic seizure by differentiation and spectral analysis of electroencephalograms. *Comput. Biol. Med.* **66**, 352–356. https://doi.org/10.1016/j.compbiomed.2015.04.034 (2015).
41. Ouyang, C. S., Yang, R. C., Wu, R. C., Chiang, C. T. & Lin, L. C. Determination of antiepileptic drugs withdrawal through EEG Hjorth parameter analysis. *Int. J. Neural Syst.* **30**, 1–16. https://doi.org/10.1142/S0129065720500367 (2020).
42. Pepi, C. *et al.* Can presurgical interhemispheric EEG connectivity predict outcome in hemispheric surgery? A brain machine learning approach. *Brain Sci.* **13**, 71. https://doi.org/10.3390/brainsci13010071 (2022).
43. Yossofzai, O. *et al.* Development and validation of machine learning models for prediction of seizure outcome after pediatric epilepsy surgery. *Epilepsia* **63**, 1956–1969. https://doi.org/10.1111/epi.17320 (2022).
44. Arle, J. E., Perrine, K., Devinsky, O. & Doyle, W. K. Neural network analysis of preoperative variables and outcome in epilepsy surgery. *J. Neurosurg.* **90**, 998–1004. https://doi.org/10.3171/jns.1999.90.6.0998 (1999).
45. Garcia Gracia, C. *et al.* Seizure freedom score: A new simple method to predict success of epilepsy surgery. *Epilepsia* **56**, 359–365. https://doi.org/10.1111/epi.12892 (2015).
46. Gracia, C. G. *et al.* Predicting seizure freedom after epilepsy surgery, a challenge in clinical practice. *Epilepsy Behav* **95**, 124–130. https://doi.org/10.1016/j.yebeh.2019.03.047 (2019).
47. Sinclair, B. *et al.* Machine learning approaches for imaging-based prognostication of the outcome of surgery for mesial temporal lobe epilepsy. *Epilepsia* **63**, 1081–1092. https://doi.org/10.1111/epi.17217 (2022).
48. Dewi, T., Masruhim, M. A. R. S. Recommending Training Set Sizes for Classification Contact. Lab Penelit Dan Pengemb FARMAKA Trop Fak Farm Univ Mualawarman, Samarinda, Kalimantan Timur 5–24 (2016).
49. Armañanzas, R. *et al.* Machine learning approach for the outcome prediction of temporal lobe epilepsy surgery. *PLoS One* **8**, e62819. https://doi.org/10.1371/journal.pone.0062819 (2013).
50. Eriksson, M. H. *et al.* Predicting seizure outcome after epilepsy surgery: Do we need more complex models, larger samples, or better data?. *Epilepsia* https://doi.org/10.1111/epi.17637 (2023).
51. Smolyansky, E. D., Hakeem, H., Ge, Z., Chen, Z. & Kwan, P. Machine learning models for decision support in epilepsy management: A critical review. *Epilepsy Behav.* https://doi.org/10.1016/j.yebeh.2021.108273 (2021).
52. Linardatos, P., Papastefanopoulos, V. & Kotsiantis, S. Explainable AI: A review of machine learning interpretability methods. *Entropy* **23**, 1–45. https://doi.org/10.3390/e23010018 (2021).
53. Ein Shoka, A. A., Dessouky, M. M., El-Sayed, A. & Hemdan, E. E. D. EEG seizure detection: Concepts, techniques, challenges, and future trends. *Multimed. Tools Appl.* https://doi.org/10.1007/s11042-023-15052-2 (2023).
54. Ayman, U. *et al.* Epileptic patient activity recognition system using extreme learning machine method. *Biomedicines* **11**, 816. https://doi.org/10.3390/biomedicines11030816 (2023).
55. Gallotto, S. & Seeck, M. EEG biomarker candidates for the identification of epilepsy. *Clin. Neurophysiol. Pract.* **8**, 32–41. https://doi.org/10.1016/j.cnp.2022.11.004 (2023).
56. Zhuang, X. *et al.* Single-scale time-dependent window-sizes in sliding-window dynamic functional connectivity analysis: A validation study. *Neuroimage* **220**, 117111. https://doi.org/10.1016/j.neuroimage.2020.117111 (2020).
57. Yotov, K., Hadzhikolev, E. & Hadzhikoleva, S. Determining the number of neurons in artificial neural networks for approximation, trained with algorithms using the Jacobi matrix. *TEM J.* **9**, 1320–1329. https://doi.org/10.18421/TEM94-02 (2020).
58. Babani, L., Jadhav, S. & Chaudhari, B. Scaled conjugate gradient based adaptive ANN control for SVM-DTC induction motor drive. *IFIP Adv. Inf. Commun. Technol.* **475**, 384–395. https://doi.org/10.1007/978-3-319-44944-9_33 (2016).

Acknowledgements

Author contributions

M.M. contributed to the conception, design, analysis of data and have drafted the article; C.P., and G.C.P., contributed to the acquisition and extraction of data, A.D.B., and M.C.R., contributed to the interpretation of data, G.P., contributed to analysis of data, F.V., and C.E.M., contributed to the revision of the article, N.S., contributed to the drafting and revision of the article, L.D.P., contributed to the conception, design, interpretation of data, and to the drafting and revision of the article. All authors reviewed and approved the submitted version of this article and agreed both to be personally accountable for the author's own contributions and to ensure that questions related to the accuracy or integrity of any part of the work, are appropriately investigated, resolved, and the resolution documented in the literature.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-60622-5.

Correspondence and requests for materials should be addressed to N.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.