British Journal of Clinical Pharmacology

Br J Clin Pharmacol (2018) **84** 2615–2624 2615

ORIGINAL ARTICLE

Personalized prediction model for seizure-free epilepsy with levetiracetam therapy: a retrospective data analysis using support vector machine

Correspondence Xiong Han, Department of Neurology, Zhengzhou University People's Hospital, Henan Provincial People's Hospital, Henan Province, China. Fax: +86 0371 6589 7586; E-mail: xionghan_doc@126.com

Received 1 November 2017; **Revised** 8 July 2018; **Accepted** 12 July 2018

Jia-hui Zhang[1,*], Xiong Han[1,*], Hong-wei Zhao[2], Di Zhao[3], Na Wang[1], Ting Zhao[1], Gui-nv He[1], Xue-rui Zhu[1], Ying Zhang[1], Jiu-yan Han[4] and Dian-ling Huang[3]

[1]Department of Neurology, Zhengzhou University People's Hospital, Henan Provincial People's Hospital, Henan Province, China, [2]Department of Pharmacy, Zhengzhou University People's Hospital, Henan Provincial People's Hospital, Henan Province, China, [3]Department of Computer Network Information Center, Chinese Academy of Sciences, Beijing, China, and [4]Department of Clinical Medicine, Zhengzhou University, Henan Province, China

*These authors contributed equally to the manuscript.

Keywords effectiveness, epilepsy, methodology, neuroscience, therapeutics

AIMS

To predict the probability of a seizure-free (SF) state in patients with epilepsy (PWEs) after treatment with levetiracetam and to identify the clinical and electroencephalographic (EEG) factors that affect outcomes.

METHODS

Retrospective analysis of PWEs treated with levetiracetam for 3 years identified 22 patients who were SF and 24 who were not. Before starting levetiracetam, 11 clinical factors and four EEG features (sample entropy of α, β, θ, δ) were identified. Overall, 80% of each the two groups were chosen to establish a support vector machine (SVM) model with 5-fold cross-validation, hold-out validation and jack-knife validation. The other 20% were used to predict the efficacy of levetiracetam. The mean impact value (MIV) algorithm was used to rank the relativity between factors and outcomes.

RESULTS

Compared with SF patients, not SF patients displayed a specific decrease in EEG sample entropy in α band from the F4 channel, β band from Fp2 and F8 channels, θ band from C3 channel ($P < 0.05$). The SVM model based on the clinical and EEG features yielded 72.2% accuracy of 5-fold cross-validation, 75.0% accuracy of jack-knife validation, 67.7% accuracy of hold-out validation in the training set and had a high prediction accuracy of 90% in test set (sensitivity was 100%, area under the receiver operating characteristic curve was 0.96). The feature of β band from Fp2 weighs heavily in the prediction model according to the mean impact value algorithm.

CONCLUSIONS

The efficacy of levetiracetam on newly diagnosed PWEs could be predicted using an SVM model, which could guide antiepileptic drug selection.

DOI:10.1111/bcp.13720

WHAT IS ALREADY KNOWN ABOUT THIS SUBJECT

- Effectiveness of the approved first-line antiepileptic drugs does not differ substantially.
- Failure of the first attempted antiepileptic drugs could be a risk factor for refractory epilepsy, but the drugs selection is still an *expert-based* clinical decision.
- The model of support vector machine can predict epilepsy surgery outcome.

WHAT THIS STUDY ADDS

- The efficacy of levetiracetam on newly diagnosed patients with epilepsy could be predicted using a support vector machine model.
- The clinical and electroencephalographic factors that affect therapy outcomes could also be identified.
- The new method of efficacy analysis in clinical pharmacology could guide antiepileptic drug selection and contribute to truly personalized therapy.

Introduction

Antiepileptic drugs (AEDs), with an effectiveness of about 60%, are the first treatment choice for most patients with epilepsy (PWEs) [1]. Significantly, the efficacy of drugs in the early stages of the disease is particularly important because early efficacy is closely related to the long-term prognosis. Thus, failure of the first attempted AED could be a risk factor for refractory epilepsy [2]. In addition, selecting an effective AED could reduce the cost, shorten the course of treatment, and relieve the pain of seizures. Currently, however, only about 40% of newly diagnosed PWEs become seizure-free after the first drug [3], so more than half of the patients must try a second or even a third AED. Even so, 30–40% of patients still suffer from the consequences of refractory epilepsy [4]. Truly personalized therapy to control seizures is still in the future.

Selecting a personalized AED for a newly diagnosed PWE is essential but challenging, particularly as the effectiveness of the approved first-line agents does not differ substantially [5]. In addition to the drug's side effects and mechanisms, when choosing an AED, doctors often need to consider numerous clinical features, including the patients' age, sex, seizure type [3, 4, 6–10] and electroencephalographic (EEG) features, such as the duration of the EEG seizure and the power spectrum of the δ band [11, 12], which have been deemed closely related to the prognosis. Therefore, we hypothesized that the efficacy of a drug could be predicted using an algorithm to analyse comprehensively the clinical and EEG features before prescribing that AED.

The support vector machine (SVM) is a machine-learning algorithm that has shown many advantages in solving classic classification and regression cases [13]. The SVM has been applied to seizure prediction [14], detection [15] and patient classification [16]. He *et al*. [17] successfully predicted recurrence preoperatively by establishing an SVM model, proving that the model could predict the probable outcome. An efficacy prediction model for AED outcomes, however, is still lacking.

Levetiracetam (LEV) is one of the most commonly used AEDs. In this study, we aimed to extract the relative clinical and EEG features of PWEs before they took LEV and then evaluate, via the SVM, the data relative to the goal of predicting seizure outcome.

Methods

The proposed algorithm consists of feature extraction, SVM classification using 5-fold cross-validation, hold-out validation, jack-knife validation and model evaluation (Figure 1).

Patient database

Forty-six newly diagnosed PWEs at the Epilepsy Center of Henan Provincial People's Hospital between 2014 and 2016 were studied retrospectively. Inclusion criteria included the following: (i) presence of an epileptic syndrome, epilepsy and/or epileptic seizures as defined according to the guidelines of the International League against Epilepsy [18, 19]; (ii) LEV accepted as the medication after taking the medical history, neurological examination, scalp EEG and magnetic resonance imaging (MRI); (iii) regular follow-up for after 1 month, 3 months, 6 months 12 months. Exclusion criteria were as follows: (i) acute symptom onset; (ii) follow-up for <1 year; (iii) seizures during pregnancy; (iv) neuropsychiatric drugs were taken before LEV; (v) poor compliance (Figure A1 in the Appendices).

The Henan Provincial People's Hospital for Research with Human Subjects approved the study (Ethical Approval 2015 Round No.13). All participants provided written informed consent.

Based on the last outpatient or telephone follow-up records, patients meeting the criteria for Engel class I [20] were classified as SF. Patients who met criteria for Engel class II, III, or IV [20] were classified as not SF (NSF).

Features extraction

Clinical features. After MRI, EEG and neurological examination, we recorded the following clinical features: (i) age; (ii) duration of epilepsy; (iii) family history of epilepsy; (iv) seizure type (generalized, focal, or unknown onset); (v) seizure frequency before LEV (mean number of partial and generalized seizures per month over the past 12 months [21]); (vi) with or without comorbidities (psychiatric disorders such as depression, anxiety disorder, psychosis [22]); (vii) seizure circadian rhythm (increased seizure occurrence during the day, night, or both [23]); (viii) presence (or not) of temporal lobe epilepsy; (ix) time between LEV initiation to the last seizure before LEV; (x) with or without interictal spikes; (xi) with or without MRI findings.

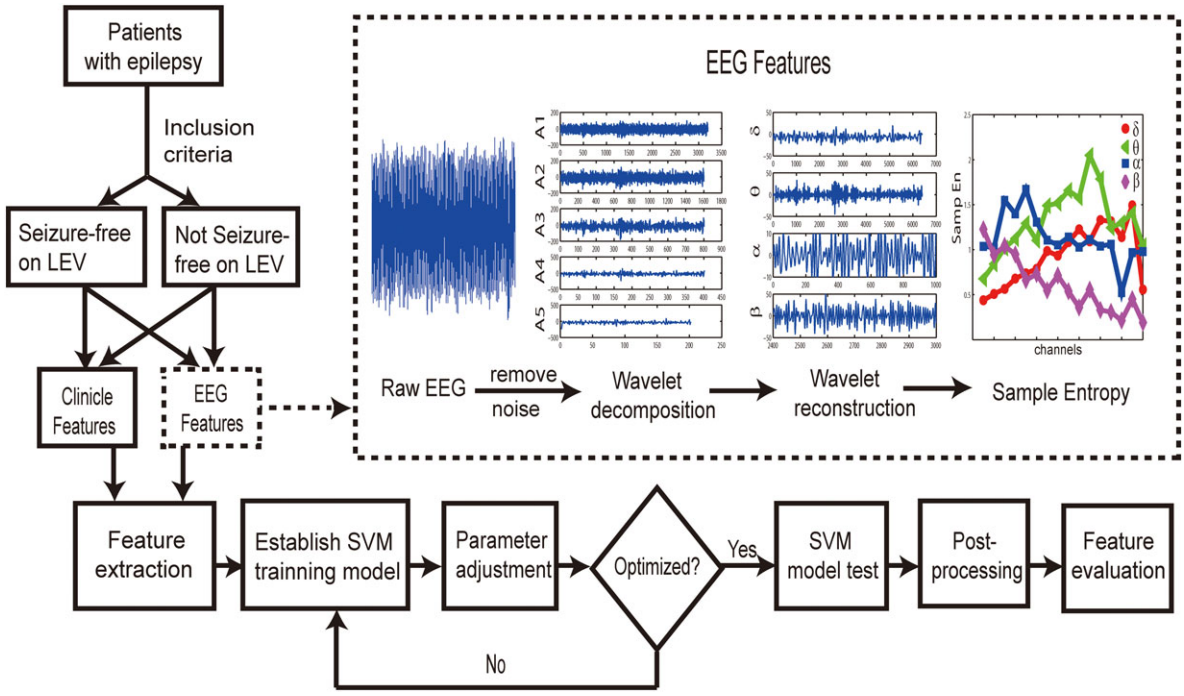

Figure 1

Architecture of the support vector machine (SVM)-based outcome prediction system. Various stages of the algorithm such as feature extraction, classification and post-processing are schematically shown. The detailed process of electroencephalography (EEG) features extraction is present in dashed box which includes converting raw EEG outputs to Sample entropy. LEV, levetiracetam

EEG features. Different EEG devices can have different parameters such as the number of electrodes, sample rate, amplifier specs and sampling time. To eliminate any bias caused by differing EEG devices and parameters, we need to use strictly uniform standards for signal acquisition to make the model more accurate. EEG recordings from the same digital EEG machine (Nation 9128 W; NCC Medical Co., Ltd, Shanghai, China) were evaluated internally. Eighteen electrodes were placed according to the international 10–20 system – specifically, Fp1, Fp2, F7, F8, F3, F4, C3, C4, T3, T4, T5, T6, P3, P4, O1, O2 – with two placed in the bilateral ears as reference electrodes. EEG signals were collected for at least 30 min and the sampling rate was 128 Hz.

For this study, we used Sample Entropy (SampEn) of δ, θ, α, β to represent the EEG features. SampEn [24] is a well-defined statistical concept used to measure complexity within dynamic processes and is also a non-linear feature of EEG, which can be calculated as follows.

1. Denote a time series of length N by $X(1),X(2),X(3)\cdots X(N)$ and construct an embedding vector with m consecutive data points:

$$X(i) = [X(i), X(i+1), \cdots, X(i+m-1)],\ \ i = 1 \sim N - m + 1.$$

2. The expression (b) $d[X(i), X(j)]$ presents the Chebyshev distance and defines for each i $(1 \le i \le N - m)$:

$$d[X(i), X(j)] = max[|X(i+k) - X(j+k)|], k = 0 \sim m - 1.$$

3. The expression (c) r specifies a tolerance value, and $r > d[X(i), X(j)]$.
4. The expression (d) $B_r^m(r)$ represents the proportion of $X(j)$ whose distance to $X(i)$ is less than r. $B^m(r) = (N - m + 1)^{-1} \cdot \sum_{i=1}^{N-m} B_r^m(r)$.
5. Similarly, for each $i \le N - m + 1$, we also define $B^{m+1}(r) = N^{m+1}(i)/(N - m - 1)$, where B^{m+1} represents the proportion corresponding to the dimension of $m + 1$.
6. SampEn is calculated as:

$$SampleEn(m, r, N) = -\ln\left[B^{m+1}(r)/B^m(r)\right], \text{where m} = \text{embedding dimension; r} = \text{threshold value; N} = \text{data length.}$$

The process of extracting SampEn includes data preprocessing, wavelet decomposition, wavelet reconstruction, and SampEn production. For one EEG, 100 s of signal (12 800 epochs) were quantitatively processed using MATLAB. Epochs were selected from EEGs taken in awake status, with eyes open and without any epileptiform discharges or obvious artefacts. Artefacts were further eliminated by independent component analysis [25]. Coefficients were obtained after three-layer wavelet decomposition, and then were used to reconstruct the bands of δ (0.5–4.0 Hz), θ (4–8 Hz), α (8–13 Hz) and β (13–30 Hz). Finally, we use SampEn to calculate the four bands in every channel for each patient.

Statistical analysis

SVM model. Finally, we tested whether clinical and EEG features could successfully predict seizure outcome using the SVM model. SVM is a classic machine-learning model

that has an advantage in solving the problem of small samples [26, 27]. Theoretically, without considering computational cost, learning speed, or other factors [28], some empirical dependence models such as the neural network model need larger samples to avoid over-fitting and under-fitting. However, for structural dependence models such as SVM, the special principle and algorithm is only to find a hyper-plane in high dimensional space with feature vectors (clinical and EEG features in this study) from the samples. For prediction, new samples are then projected into the same space and assigned to belong to a class based on which side of the margin they fall. It means that when the data and feature distribution are good, increasing training data will not have any impact on the classification because of the perfect hyper-plane. Therefore, we used the SVM as the efficacy prediction model to solve the limited sample.

The performance of SVM relies upon the kernel selection. Radial basis function (RBF) was used as it is the most commonly kernel function used to map data into a space [29]. The general modelling process was to divide the samples randomly into training and test sets to check the generalization of the model. The 80% : 20% ratio for training:prediction for an SVM model can ensure suitable parameters.

To prevent interference from invalid EEG features, only the band that had significant differences in the two groups were extracted to represent the EEG features. In consideration of the actual AEDs selection in clinic, every clinical feature in the study was also regarded as important factors in model. So we tested three models separately for their: (i) selected EEG features; (ii) clinical features only; and (iii) variable sets of (i) and (ii) combined. In our algorithm, the binary classification is performed in two steps: establishing the model and examining its function.

For the first step, we randomly selected about 80% (total 36 patients) of patients from each of the SF and NSF groups as training sets to build the model. The exacted features from these patients were conducted by a Lib-SVM classifier [30]. A Lib-SVM model requires two parameters: a *kernel* and a *regularization* parameter (Cost and γ). Cost (C parameter) can control the smoothness of the decision boundary in the transformed space. The γ parameter is set in kernel function and determining the distribution of data mapped to a new feature space.

In this study, we used an RBF kernel [$k(x, x') = \exp(\gamma|x\text{-}x'|^2)$], and the regularization parameter (Cost and γ) was identified using a *grid search* method within a 5-fold cross-validation procedure. For this procedure, the data set was randomly divided into five subsets, with each subset used as a *validation set* and the remaining four used as the training set to create the model. This process was repeated five times such that every portion was used to assess the performance of the models. The tunable model parameters were iterated to minimize absolute mean errors, i.e. differences between predicted and measured output values on the validation set. The performance of the models was the average scores of the model trained on each fold. The tuning model parameters were optimized from the training data using the cross validation method.

At present, besides the k-fold validation, the most frequently used parameters to evaluate the generalization are hold-out validation and resampling validation (bootstrapping and jack-knife). In the jack-knife validation, each patient is singled out in turn as a test set and the remaining patients are used as training set. In the hold-out validation, the *validation set* was equally divided into two subsets, with each subset used as a test set and the remaining one used as the training set to create the model. So we also used the methods of jack-knife validation, hold-out validation to show the feasibility of the model further in classifying and predicting the two groups of patients.

In the second step, the remaining about 20% (totally 10 patients) of patients in each group were utilized as a test set to validate the function. The prediction performance of the SVM was evaluated using the statistical measures of accuracy ($= \frac{TP+TN}{(TP+TN+FP+FN)}$), sensitivity ($= \frac{TP}{(TP+FN)}$), specificity ($= \frac{TN}{(TN+FP)}$), positive predictive value ($= \frac{TP}{(TP+FP)}$) and negative predictive value ($= \frac{TN}{(TN+FN)}$), where true positive (TP) is the number of segments recognized as NSF both by the algorithm and by the neurologist; false positive (FP) denotes the number of segments differentiated as SF by classifier but their true labels are NSF; true negative (TN) is the number of patients classified as SF both by the algorithm and by the EEG experts; false negative (FN) is the number of SF patients misclassified as NSF by the algorithm. Additionally, a receiver operating characteristic curve and the area under the receiver operating characteristic curve (AUC) were generated.

Feature evaluation. The mean impact value (MIV) in this study is an important index used to select the independent features that have a great impact on the model. To date, it has been considered one of the best indexes to assess assigned coefficient values [31]. In this study, after training the SVM model, two new training sets were obtained each time an independent feature increased or decreased 10% that were used for simulation according to the fitting model. Thus, the mean of the difference in the features of the two simulation results was calculated according to the number of samples (i.e. the MIV). Finally, the sequence of the features was sorted according to their absolute MIVs, and the related compounds were identified as the potential lead constituents. Although the selected features may not have significance determined by a t-test, they can have a classification value in the model.

Statistical methods. SPSS software was used for statistical data analysis. The clinical data and the SampEn of the EEG signals were compared between the SF and NSF groups. Mean ± standard deviation together with independent sample *t* tests were used to describe and compare quantitative data with a normal distribution, and the Mann–Whitney *U* test was applied for abnormal distributions, as appropriate. The χ^2 test was used for qualitative data. For all measures, a value of $P < 0.05$ was considered to indicate statistical significance. No adjustment of the α-level for multiplicity has been made.

Results

Clinical characteristics

A total of 46 PWEs were enrolled, 22 of whom showed remission of their seizures (SF group), and the other 24 PWEs were in the NSF group. Maintenance dose of LEV in SF group was

less than NSF group ($P = 0.004$), but there were no significant differences in the 11 selected clinical characteristics used in SVM between the SF and NSF groups ($P \geq 0.05$; Table 1).

EEG features

The SF and NSF groups showed no differences in the Fp1, C4, F3, P3, P4, O1, O2, T3, T4, T5, T6 regions ($P \geq 0.05$). Compared with the NSF group, however, the SF group produced significantly higher SampEN in the α band from the F4 channel ($P < 0.001$), β band from the Fp2 channel ($P = 0.008$), β band from the F8 channel ($P = 0.011$) and θ band from the C3 channel ($P < 0.001$; Figure 2).

Personalized prediction model for SF

The lib-SVM model with RBF kernel ($\gamma = 0.04$ and Cost = 16.0) using clinical and EEG features successfully predicted the efficacy of LEV with a 91.7% accuracy, a 72.2% 5-fold cross-validation, 75.0% accuracy of jack-knife validation, 67.7% accuracy of hold-out validation and a 0.95 (AUC) in the training set. Importantly, it could also predict the efficacy in 10 cases that were not used to train the model. Specifically, drug efficacy for the test data was predicted with 90.0% accuracy and had a 0.96 AUC. Thus, the success of the model was not because of a bias in sample ratios. The cross-validation and AUC results of the individual models (EEG: cross-validations, 75%; AUC, 0.84; clinical features: cross-validation, 63.9%; AUC, 0.72) indicated that the model established using only EEG features was more generalizable than that established by clinical features, but that neither was as good as the combined model. This means that the combined model is more suitable for AED selection (Table 2). The predictions for the training and test sets are listed in Table 3.

Feature evaluation

After terminating the MIV algorithm, we chose features with higher absolute MIV values as the features potentially affecting the efficacy of LEV. Age, seizure type, interictal spikes and seizure frequency before LEV did not have a significant effect on the model. The detailed sort exercise is shown in Table 4.

Table 1

Sample demographic and clinical characteristic

Sample group	SF (*n* = 22)	NSF (*n* = 24)
Age, mean ± SD, years	19.5 ± 9.5	24.0 ± 12.1
Age at epilepsy onset, mean ± SD, years	15.2 ± 6.2	20.5 ± 13.0
Duration of epilepsy, mean ± SD, years	4.4 ± 6.1	3.8 ± 5.6
Interictal spike (Y/N), *n*	12/10	13/11
Seizure frequency before LEV, times/month	3.0 ± 5.9	3.4 ± 4.3
MRI findings (Y/N), *n*	6/16	12/12
Seizure circadian rhythm (day/night/both), *n*	8/8/6	8/5/11
Temporal lobe epilepsy (Y/N), *n*	5/17	11/13
Comorbidity (Y/N), *n*	8/14	7/17
Family history (Y/N), *n*	2/20	0/24
LEV initiation to the last seizure, days	6.5 ± 5.3	8.1 ± 7.8
Seizure type, *n*		
Focal	10	7
Generalized	11	15
Combined two types	1	2
Maintenance dose of LEV, mg BID*	568.2 ± 290.5	760.4 ± 260.4
Follow-up time, month	20.2 ± 6.5	20.2 ± 6.7

BID, twice daily; LEV, levetiracetam; MRI, magnetic resonance imaging; N, No; NSF, Not seizure-free group; SD, standard deviation; SF, Seizure-free group; Y, Yes

**vs.* control, $P < 0.05$. For continuous variables, independent-sample *t* tests or Mann–Whitney *U* test was carried out. For categorical variables, χ^2 tests were carried out

Discussion

This study represents our first attempt to use the SVM algorithm to predict whether PWEs could achieve an SF state by taking LEV, which could provide guidance for the selection of LEV in newly diagnosed PWEs.

LEV is an effective AED and has been widely used in various patient groups, such as those with seizures, syndromes, or refractory epilepsy. Perry and Benatar [32] had demonstrated that 57% of epileptic children aged <4 years could achieve seizure remission after LEV treatment. Wu *et al.* [21] found that, for adult patients, the responder rate (patients with ≥50% reduction in seizure frequency) was 68% and the control rate (for SF patients) was 39%. Ben-Menachem and Falter's study [33], with a high dose of LEV (3000 mg day^{-1}) to treat refractory epilepsy proved that the responder rate of this LEV monotherapy was 59.2% and the control rate 18.4%, whereas with polytherapy the responder rate was 42.1% and the control rate 8.2%. Berkovic *et al.*'s study [34] on idiopathic generalized epilepsy found that 72.2% of the subjects responded, and 34.2% achieved an SF state. In summary, when selecting LEV by depending on clinical experience alone, the responder rate was 42.1–72.2% and the control rate only 8.2–57.0%. Hence, this model offers prediction strength of 90% accuracy regarding the outcome, which could improve the efficacy of LEV and avoid refractory epilepsy.

The model used to predict the efficacy of an AED in PWEs in this study conforms to the principle of SVM. First, SVM is a machine-learning model that can be used to predict efficiency. Vidyasagar [35] proposed that SVM can extract a small number of features among tens of thousands of measured features and then accurately predict a tumour's response to the drug. He *et al.* [17] established an SVM model through preoperative functional MRI and successfully predicted the recurrence rate of temporal lobe epilepsy: the 76% rate of prediction accuracy they reported was similar to that achieved with *expert-based* clinical decisions. Colic *et al.* [12] established an SVM model through the EEG δ band in mice to predict the efficacy of midazolam for treating Rett syndrome. Their predictive rate was 77%, which proved that the efficacy of AED could be predicted by SVM.

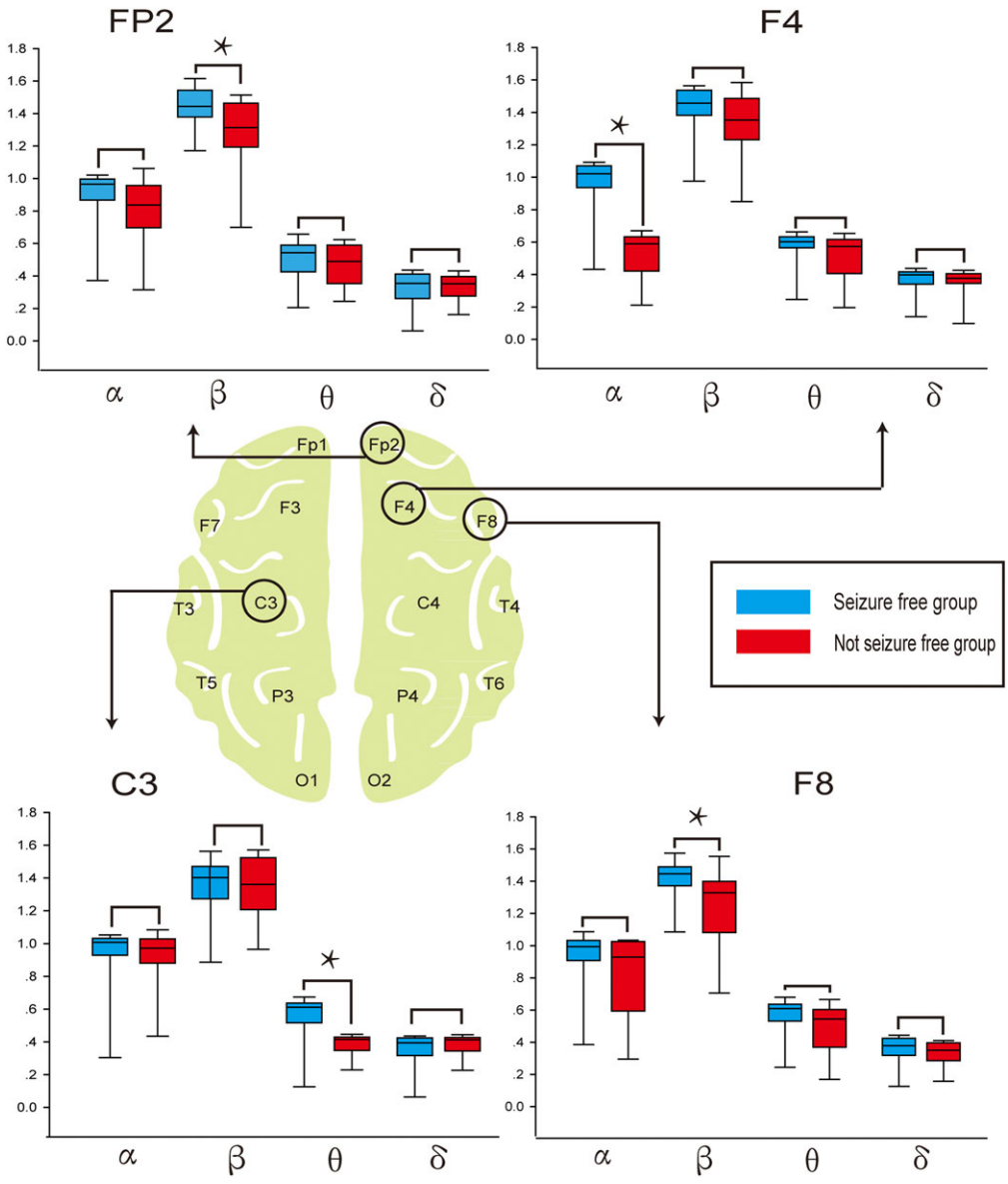

Figure 2

Sample entropy of the bands with a significant difference between seizure-free (SF) and not seizure-free (NSF) groups. Graphic presentation (boxplot diagrams) of relative sample entropy within each frequency band in channels FP2, F4, C3, and F8, which are significantly different between the SF and NSF patients. The lower and upper borders of the rectangular box correspond to the 25% and 75% percentiles of the data, with the median indicated by the black line. The red box represents the SF group, and the blue box represents the NSF group. *indicate statistically significant results. Compared with SF patients, the NSF patients had significantly decreased Sample Entropy in the β frequency band of Fp2, α frequency band of F4, θ frequency band of C3, and β frequency band of F8 (Mann–Whitney *U* test, $P < 0.05$)

Second, the features extracted in this study are closely related to the efficacy of the AED. Lamberink *et al.*'s study [36], which focused on seizure recurrence after withdrawal of AEDs, affirmed that the duration of epilepsy, seizure frequency, family history, seizure type and interictal spikes were factors affecting the efficacy. It was also determined that age [7], temporal lobe epilepsy [9], comorbidities [10], seizure circadian rhythm [23] and MRI findings [37] might be related to the prognosis. Therefore, the good performance of the model established by these features is reasonable.

Finally, validation—assessment of how well a prediction works on data other than those on which the model was built—is deemed the most important issue in prognostic modelling [38] In this study, a 5-fold cross-validation method was used to verify the generalization of the classifier, the results of which are similar to those of He *et al.* [17], indicating that it is feasible that the model could predict the efficacy of LEV.

Another observation is the effects of various features on the efficacy calculated by the MIV. The SampEn of the β band from Fp2 channel contributes most to the model, which,

Table 2

Treatment outcome prediction summary

	ACC	SEN	SPE	PPV	NPV	ROC-AUC	Cross-validations	Jack- knife	Hold-out Validation
EEG + Clinical characteristics									
Training (*n* = 36)	91.7%	90.0%	93.8%	94.7%	88.2%	0.95	72.2%	75.0%	67.7%
Test (*n* = 10)	90.0%	100%	80.0%	83.3%	100%	0.96			
EEG characteristics									
Training (*n* = 36)	77.8%	79.0%	76.5%	79.0%	76.5%	0.89	75.0%	76.5%	70.6%
Test (*n* = 10)	70.0%	80.0%	60.0%	66.7%	75.0%	0.84			
Clinical characteristics									
Training (*n* = 36)	80.6%	83.3%	77.8%	79.0%	82.4%	0.83	63.9%	58.3%	61.1%
Test (*n* = 10)	70.0%	80.0%	60.0%	66.7%	75.0%	0.72			

ACC, accuracy; EEG, electroencephalography; NPV, negative predictive value; PPV, positive predictive value; ROC-AUC, area under the receiver-operating-characteristic curve; SEN, sensitivity; SPE, specificity

Table 3

Treatment outcome prediction summary

Model		TP	FP	TN	FN	All (NSF/SF)
EEG + Clinical characteristics	Training set	18	1	15	2	36 (19/17)
	Test set	5	1	4	0	10 (5/5)
EEG characteristics	Training set	15	4	13	4	36 (19/17)
	Test set	4	2	3	1	10 (5/5)
Clinical characteristics	Training set	15	4	14	3	36 (19/17)
	Test set	4	2	3	1	10 (5/5)

EEG, electroencephalography; FN (false negative), the number of SF patients who were misclassified as NSF by the algorithm; FP (false positive), the number of patients who were categorized as SF by the classifier but were actually NSF; NSF, Not seizure-free group; SF, seizure-free group; TN (true negative), the number of patients classified as SF by both the algorithm and the neurologist; TP (true positive), the number of patients who were recognized as NSF by both support vector machine and the neurologist

together with other EEG features, proves that the background rhythm of interictal EEG, instead of interictal discharges, has predictive value regarding the efficacy of LEV. In fact, the prognosis of interictal discharges is still debatable. EEG abnormalities were significantly associated with outcomes [39]; in the absence of other predictive factors, however, they only slightly increased the risks [36]. In contrast, rhythm has been thought to be related to outcome. For example, the prognosis of generalized slow waves after trauma and subarachnoid haemorrhage is often poor [40]. The δ band can also serve as a predictor to evaluate the efficacy of midazolam in the mouse SVM model [12]. Moreover, the presence of a focal β frequency discharge is considered highly predictive of excellent postsurgical seizure control [41], although there is no clear explanation of the relation between the specific region or band and efficacy. Further studies are therefore needed to explore these phenomena.

EEG features weigh heavily on the outcome, but the results of the three models in the study show that it is inaccurate to predict the prognosis by EEG alone (i.e. without considering clinical features). Many clinical features are closely related to prognosis, but their degree of influence had not been studied previously. The MIV algorithm in this study showed that the outcome could vary because of changes in the identified features, especially MRI findings, family history and seizure circadian rhythm. Although age, seizure type, interictal spikes and seizure frequency before LEV had no influence on the outcome, the reasonable explanation is that LEV may have an efficacy that is similar in patients with different ages, seizure types, seizure frequencies, or with or without interictal spikes, which has been confirmed in previous research [42–44].

Our study has several limitations. First, the limited amount of data could bias their representativeness. Although fixed standards do not yet exist for the number of samples needed for SVM models, measures to improve the model will be rich by processing large samples, such as optimizing parameters, kernel function or even exchanging to other machine learning models with higher complexity. However, the sample size in medical researches is often limited and the sample size in other SVM model research on disease classification, prediction etc. is also not always large. Additionally, the establishment of the model success is not only related to the sample size, but also to the kernel function, dimension. Optimizing the SVM model is the current focus in the field of algorithms, so a model with more valuable features of SF/NSF samples and a faster processing time and higher accuracy could be established in another prospective study. Furthermore, the follow-up time of this study may be too short

Table 4

Input variables and sorting of mean impact values (MIVs)

Sequence	Features	MIV
1	F2 - β	0.1111
2	F4 - α	0.0833
3	C3 - θ	0.0556
4	MRI findings	0.0556
5	Family history	0.0556
6	Seizure circadian rhythm	0.0556
7	F8 - β	0.0278
8	LEV initiation to the last seizure, days	0.0278
9	Comorbidity	0.0278
10	Duration of epilepsy	0.0278
11	Temporal lobe epilepsy	0.1111
12	Age	0
13	Seizure type	0
14	Interictal spike	0
15	Seizure frequency before LEV	0

F2 - β, β bands from F2 region according to the International 10–20 system. Correspondingly, F4 - α, α bands from F4 region. C3 - θ, θ bands from C3 region. F8 - β, β bands from F8 region; LEV, levetiracetam; MRI, magnetic resonance imaging

to represent the final outcome of therapy. Thus, a multicentre, prospective study to expand the sample size is still necessary.

Competing Interests

There are no competing interests to declare.

We would like to thank Si-yu Zhao (Central South University) for assistance with SVM model, quantitative EEG analysis and Zhi-wei Xu (Zhengzhou University) for data statistics and Zi-qian Zhu (Zhengzhou University) for searching literatures. The author also thank Nancy Schatken BS, MT (ASCP), from Liwen Bianji, Edanz Group China (www.liwenbianji.cn/ac), for editing the English text of a draft of this manuscript. Finally, the current study was funded by Foundation of Wu Jieping (OR 320.6750.14128) and Basic and Frontier technological Project in Henan Province 2015 (OR 152300410155).

Contributors

J.–h.Z. and X.H. are joint first authors. J.–h.Z., X.H. designed the study, X.H. obtained funding. H.–w.Z., G.–n.H., J.y.H. and Y.Z. acquired the data, J.–h.Z., X.H. analyzed and interpreted the data, D.Z., N.W., T.Z. and D.–l.H. contributed to the interpretation of the results and critical revision of the manuscript for important intellectual content, J.–h.Z. and X.H. drafted the manuscript. All authors have read and approved the final manuscript.

Appendix A

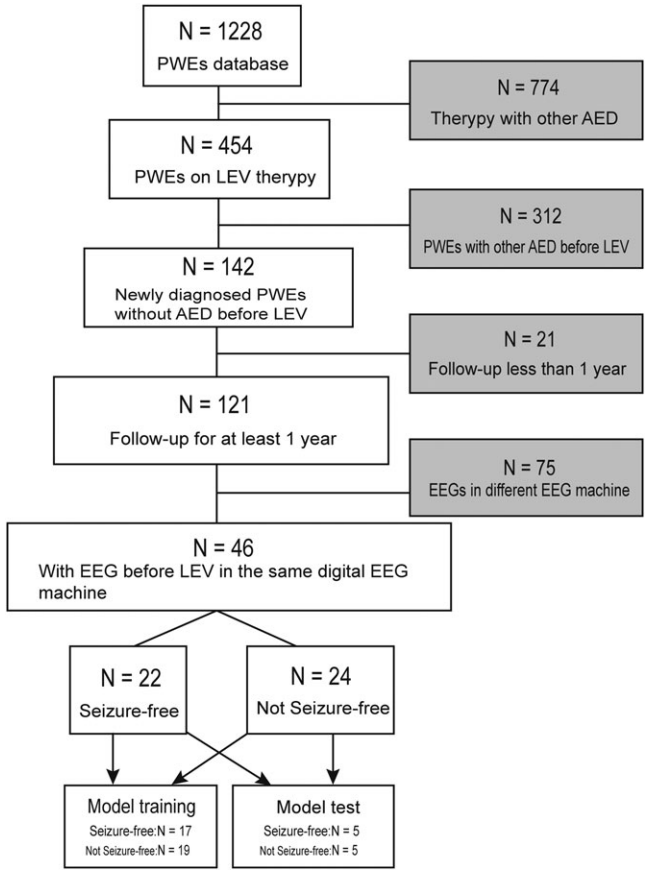

Figure A1

Diagram of patients across the study period. PWEs, patients with epilepsy; AED, antiepileptic drug.

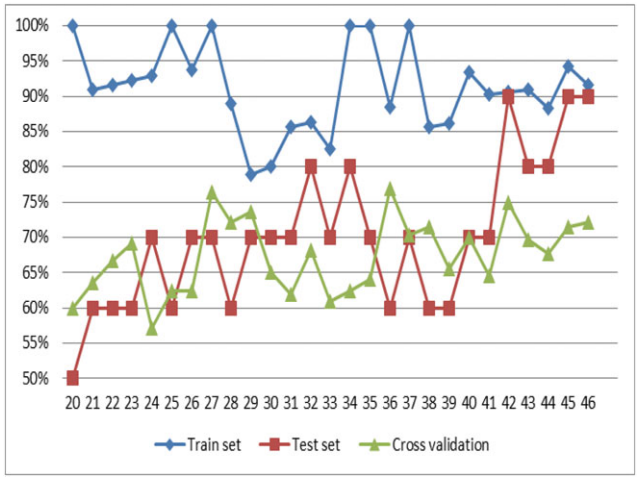

Figure A2

The result of prediction outcome with decrease of sample size. Blue line = training set; red line = test set; green line = 5-fold cross validation. With the decrease of sample size, the gap of accuracy between the training set and the test set is obvious widening, the result of 5-fold cross-validation become lower and lower.

Table A1

The result of prediction outcome with decrease of sample size

Sample size	36	35	34	33	32	31	30	29	28	27	26	25	24	23
Train set, %	91.7	94.3	88.2	90.9	90.6	90.3	93.3	86.2	85.7	100	88.5	100	100	82.6
Test set, %	90	90	80	80	90	70	70	60	60	70	60	70	80	70
Cross-validation, %	72.2	71.4	67.7	69.7	75	64.5	70	65.5	71.4	70.4	76.9	64	62.5	60.9
Sample size	**22**	**21**	**20**	**19**	**18**	**17**	**16**	**15**	**14**	**13**	**12**	**11**	**10**	**...**
Train set, %	86.4	85.7	80	79.0	88.9	100	93.8	100	92.9	92.3	91.7	90.9	100	
Test set, %	80	70	70	70	60	70	70	60	70	60	60	60	50	
Cross-validation, %	68.2	61.9	65	73.7	72.2	76.5	62.5	62.5	57.1	69.2	66.7	63.6	60	

References

1 Glauser T, Ben-Menachem E, Bourgeois B, Cnaan A, Guerreiro C, Kalviainen R, *et al*. Updated ILAE evidence review of antiepileptic drug efficacy and effectiveness as initial monotherapy for epileptic seizures and syndromes. Epilepsia 2013; 54: 551–63.

2 Dlugos DJ, Sammel MD, Strom BL, Farrar JT. Response to first drug trial predicts outcome in childhood temporal lobe epilepsy. Neurology 2001; 57: 2259–64.

3 Perucca E, Tomson T. The pharmacological treatment of epilepsy in adults. The Lancet Neurology 2011; 10: 446–56.

4 Ben-Menachem E. Medical management of refractory epilepsy – practical treatment with novel antiepileptic drugs. Epilepsia 2014; 55 (Suppl. 1): 3–8.

5 Kwan P, Brodie MJ. Clinical trials of antiepileptic medications in newly diagnosed patients with epilepsy. Neurology 2003; 60 (11 Suppl. 4): S2–12.

6 Bianchin MM, Velasco TR, Martins APP, Sakamoto AC. Sex as a prognostic factor for surgical outcome in mesial temporal lobe epilepsy. Arch Neurol 2007; 64: 288.

7 Ferlazzo E, Sueri C, Gasparini S, Aguglia U. Challenges in the pharmacological management of epilepsy and its causes in the elderly. Pharmacol Res 2016; 106: 21–6.

8 Zhuo C, Jiang R, Li G, Shao M, Chen C, Chen G, *et al*. Efficacy and tolerability of second and third generation anti-epileptic drugs in refractory epilepsy: a network meta-analysis. Sci Rep 2017; 7: 2535.

9 Aull-Watschinger S, Pataraia E, Czech T, Baumgartner C. Outcome predictors for surgical treatment of temporal lobe epilepsy with hippocampal sclerosis. Epilepsia 2008; 49: 1308–16.

10 de Araujo Filho GM, Yacubian EM. Juvenile myoclonic epilepsy: psychiatric comorbidity and impact on outcome. Epilepsy Behav 2013; 28 (Suppl. 1): S74–80.

11 Dlugos D, Shinnar S, Cnaan A, Hu F, Moshe S, Mizrahi E, *et al*. Pretreatment EEG in childhood absence epilepsy: associations with attention and treatment outcome. Neurology 2013; 81: 150–6.

12 Colic S, Wither RG, Min L, Zhang L, Eubanks JH, Bardakjian BL. Support vector machines using EEG features of cross-frequency coupling can predict treatment outcome in Mecp2-deficient mice. Conf Proc IEEE Eng Med Biol Soc 2015; 2015: 5606–9.

13 Cherkassky VS, Mulier F. Learning from data: concepts, theory, and methods. Dent Tech 2015; 43: 105–6.

14 Park Y, Luo L, Parhi KK, Netoff T. Seizure prediction with spectral power of EEG using cost-sensitive support vector machines. Epilepsia 2011; 52: 1761–70.

15 Zhang T, Chen W. LMD based features for the automatic seizure detection of EEG signals using SVM. IEEE Trans Neural Syst Rehabil Eng 2017; 25: 1100–8.

16 Amarreh I, Meyerand ME, Stafstrom C, Hermann BP, Birn RM. Individual classification of children with epilepsy using support vector machine with multiple indices of diffusion tensor imaging. NeuroImage Clin 2014; 4: 757–64.

17 He X, Doucet GE, Pustina D, Sperling MR, Sharan AD, Tracy JI. Presurgical thalamic "hubness" predicts surgical outcome in temporal lobe epilepsy. Neurology 2017; 88: 2285–93.

18 Scheffer IE, Berkovic S, Capovilla G, Connolly MB, French J, Guilhoto L, *et al*. ILAE classification of the epilepsies: position paper of the ILAE Commission for Classification and Terminology. Epilepsia 2017; 58: 512–21.

19 Fisher RS, Cross JH, French JA, Higurashi N, Hirsch E, Jansen FE, *et al*. Operational classification of seizure types by the international league against epilepsy: position paper of the ILAE Commission for Classification and Terminology. Epilepsia 2017; 58: 522–30.

20 Engel J Jr, Wiebe S, French J, Sperling M, Williamson P, Spencer D, *et al*. Practice parameter: temporal lobe and localized neocortical resections for epilepsy: report of the quality standards Subcommittee of the American Academy of neurology, in association with the American Epilepsy Society and the American Association of Neurological Surgeons. Neurology 2003; 60: 538–47.

21 Wu T, Chen CC, Chen TC, Tseng YF, Chiang CB, Hung CC, *et al*. Clinical efficacy and cognitive and neuropsychological effects of levetiracetam in epilepsy: an open-label multicenter study. Epilepsy Behav: E&B 2009; 16: 468–74.

22 Sander JW. Comorbidity and premature mortality in epilepsy. The Lancet 2013; 382: 1618–9.

23 van Campen JS, Valentijn FA, Jansen FE, Joels M, Braun KP. Seizure occurrence and the circadian rhythm of cortisol: a systematic review. Epilepsy Behav: E&B 2015; 47: 132–7.

24 Jia Y, Gu H, Luo Q. Sample entropy reveals an age-related reduction in the complexity of dynamic brain. Sci Rep 2017; 7: 7990.

25 McKeown M. Independent component analysis of functional MRI: what is signal and what is noise? Curr Opin Neurobiol 2003; 13: 620–9.

26 Way TW, Sahiner B, Hadjiiski LM, Chan HP. Effect of finite sample size on feature selection and classification: a simulation study. Med Phys 2010; 37: 907–20.

27 Jiang H, Ching WK, Cheung WS, Hou W, Yin H. Hadamard kernel SVM with applications for breast cancer outcome predictions. BMC Syst Biol 2017; 11 (Suppl. 7): 138.

28 Bottou LE, Cun YL. Large scale online learning. Adv Neural Inf Proces Syst 2003; 11: 1109–35.

29 Gromski PS, Xu Y, Correa E, Ellis DI, Turner ML, Goodacre R. A comparative investigation of modern feature selection and classification approaches for the analysis of mass spectrometry data. Anal Chim Acta 2014; 829: 1–8.

30 Chang CC, Lin CJ. LIBSVM: a library for support vector machines. ACM Trans Intell Syst Technol 2011; 2: 1–27.

31 Jiang JL, Su X, Zhang H, Zhang XH, Yuan YJ. A novel approach to active compounds identification based on support vector regression model and mean impact value. Chem Biol Drug Des 2013; 81: 650–7.

32 Perry MS, Benatar M. Efficacy and tolerability of levetiracetam in children younger than 4 years: a retrospective review. Epilepsia 2007; 48: 1123–7.

33 Ben-Menachem E, Falter U. Efficacy and tolerability of levetiracetam 3000 mg/d in patients with refractory partial seizures: a multicenter, double-blind, responder-selected study evaluating monotherapy. European Levetiracetam study group. Epilepsia 2000; 41: 1276–83.

34 Berkovic SF, Knowlton RC, Leroy RF, Schiemann J, Falter U. Placebo-controlled study of levetiracetam in idiopathic generalized epilepsy. Neurology 2007; 69: 1751–60.

35 Vidyasagar M. Identifying predictive features in drug response using machine learning: opportunities and challenges. Annu Rev Pharmacol Toxicol 2015; 55: 15–34.

36 Lamberink HJ, Otte WM, Geerts AT, Pavlovic M, Ramos-Lizana J, Marson AG, *et al*. Individualised prediction model of seizure recurrence and long-term outcomes after withdrawal of antiepileptic drugs in seizure-free patients: a systematic review and individual participant data meta-analysis. Lancet Neurol 2017; 16: 523–31.

37 RamachandranNair R, Otsubo H, Shroff MM, Ochi A, Weiss SK, Rutka JT, *et al*. MEG predicts outcome following surgery for intractable epilepsy in children with normal or nonfocal MRI findings. Epilepsia 2007; 48: 149–57.

38 Steyerberg EW, Moons KG, van der Windt DA, Hayden JA, Perel P, Schroter S, *et al*. Prognosis research strategy (PROGRESS) 3: prognostic model research. PLoS Med 2013; 10: e1001381.

39 Morita ME, Yasuda CL, Betting LE, Pacagnella D, Conz L, Barbosa PH, *et al*. MRI and EEG as long-term seizure outcome predictors in familial mesial temporal lobe epilepsy. Neurology 2012; 79: 2349–54.

40 Steinbaugh LA, Lindsell CJ, Shutter LA, Szaflarski JP. Initial EEG predicts outcomes in a trial of levetiracetam vs. fosphenytoin for seizure prevention. Epilepsy Behav: E&B 2012; 23: 280–4.

41 Worrell GA, So EL, Kazemi J, O'Brien TJ, Mosewich RK, Cascino GD, *et al*. Focal ictal beta discharge on scalp EEG predicts excellent outcome of frontal lobe epilepsy surgery. Epilepsia 2002; 43: 277–82.

42 Specchio N, Boero G, Michelucci R, Gambardella A, Giallonardo AT, Fattouch J, *et al*. Effects of levetiracetam on EEG abnormalities in juvenile myoclonic epilepsy. Epilepsia 2008; 49: 663–9.

43 Lyseng-Williamson KA. Levetiracetam: a review of its use in epilepsy. Drugs 2011; 71: 489.

44 Goldberg-Stern H, Feldman L, Eidlitz-Markus T, Kramer U, Perez S, Pollak L, *et al*. Levetiracetam in children, adolescents and young adults with intractable epilepsy: efficacy, tolerability and effect on electroencephalogram--a pilot study. Eur J Paediatr Neurol 2013; 17: 248–53.