International Journal of Neural Systems, Vol. 24, No. 7 (2014) 1450023 (16 pages)

DOI: 10.1142/S0129065714500233

World Scientific
www.worldscientific.com

EARLY PREDICTION OF MEDICATION REFRACTORINESS IN CHILDREN WITH IDIOPATHIC EPILEPSY BASED ON SCALP EEG ANALYSIS

LUNG-CHANG LIN
Department of Pediatrics, School of Medicine
College of Medicine, Kaohsiung Medical University
Departments of Pediatrics, Kaohsiung Medical University Hospital
Kaohsiung Medical University
No. 100, Shih-Chuan 1st Rd.
Kaohsiung City 80708, Taiwan

CHEN-SEN OUYANG*
Department of Information Engineering, I-Shou University
No. 1, Sec. 1, Syuecheng Rd., Dashu District
Kaohsiung City 84001, Taiwan
ouyangcs@isu.edu.tw

CHING-TAI CHIANG
Department of Computer and Communication
National Pingtung Institute of Commerce
No. 51, Min-Sheng E. Rd., Pingtung 90004, Taiwan

REI-CHENG YANG
Departments of Pediatrics, Kaohsiung Medical University Hospital
Kaohsiung Medical University, No. 100, Shih-Chuan 1st Rd.
Kaohsiung City 80708, Taiwan

RONG-CHING WU
Department of Electrical Engineering, I-Shou University
No. 1, Sec. 1, Syuecheng Rd., Dashu District
Kaohsiung City 84001, Taiwan

HUI-CHUAN WU
Departments of Pediatrics, Kaohsiung Medical University Hospital
Kaohsiung Medical University, No. 100, Shih-Chuan 1st Rd.
Kaohsiung City 80708, Taiwan

Accepted 29 May 2014
Published Online 8 August 2014

Refractory epilepsy often has deleterious effects on an individual's health and quality of life. Early identification of patients whose seizures are refractory to antiepileptic drugs is important in considering the use of alternative treatments. Although idiopathic epilepsy is regarded as having a significantly lower risk factor of developing refractory epilepsy, still a subset of patients with idiopathic epilepsy might be refractory to medical treatment. In this study, we developed an effective method to predict the refractoriness of idiopathic epilepsy. Sixteen EEG segments from 12 well-controlled patients and 14 EEG segments from 11 refractory patients were analyzed at the time of first EEG recordings before antiepileptic drug treatment. Ten crucial EEG feature descriptors were selected for classification. Three

*Corresponding author.

of 10 were related to decorrelation time, and four of 10 were related to relative power of delta/gamma. There were significantly higher values in these seven feature descriptors in the well-controlled group as compared to the refractory group. On the contrary, the remaining three feature descriptors related to spectral edge frequency, kurtosis, and energy of wavelet coefficients demonstrated significantly lower values in the well-controlled group as compared to the refractory group. The analyses yielded a weighted precision rate of 94.2%, and a 93.3% recall rate. Therefore, the developed method is a useful tool in identifying the possibility of developing refractory epilepsy in patients with idiopathic epilepsy.

Keywords: Early prediction; refractory seizure; EEG; idiopathic epilepsy; feature descriptor; support vector machine.

1. Introduction

Epilepsy is the most common chronic disease in pediatric neurology. About 0.5–1% of children develop epilepsy during their lifetime.[1] Although 60–70% of patients are controlled by antiepileptic drugs (AED), still 30–40% of epileptic children are not controllable by AED,[2] and are called refractory epilepsy. Regardless of etiology, children with refractory epilepsy are exposed to a variety of physical, psychological, and social morbidities.[3] Patients whose seizures are difficult to control could benefit from nonpharmacological therapies, including epilepsy surgery, deep brain stimulation,[4] or ketogenic diets.[5,6] Therefore, the early identification of patients whose seizures are refractory to AED would allow them to receive alternative therapies at a more appropriate time.

Some studies report several risk factors for developing refractory epilepsy, including epileptic syndromes, age of onset, etiology, developmental status, electroencephalography (EEG) features, neuroimaging findings, and seizure frequencies before treatment.[7,8] Age at onset of epilepsy under 1 year, remote symptomatic etiology, developmental delay/mental retardation, abnormal EEG background, frequent epileptiform discharges abnormal neuroimaging, a multiple first seizure, West syndrome, more than one recurrence during the first six months after diagnosis and more than five seizures before diagnosis, are all significant predictors of a higher risk for developing refractory epilepsy. Conversely, idiopathic etiology is a significant predictor of a lower risk.[2,7,8] However, still a subset of patients with idiopathic epilepsy might be refractory to medical treatment.[9,10] EEG analysis is widely employed to investigate brain disorders and to study brain electrical activity.[11] Although EEG experts are able to qualitatively identify specific EEG abnormalities by visual inspection, some EEG parameters, including power, frequencies, and complexity cannot be analyzed without computer technology.[12] In addition, visual screening relies on the EEG experts which can also result in subjective conclusions at times. Therefore, many approaches of computer-aided EEG analysis have been developed and are applied to diagnose and treat neurological diseases, such as epilepsy,[11,13–25] Alzheimer disease,[26–31] attention-deficit hyperactivity disorder,[32–34] autistic spectrum disorder,[35,36] and major depressive disorder.[37] Among these approaches, several linear and nonlinear features have been proposed or employed for EEG analysis, such as correlation dimension (CD),[18,30,38] largest Lyapunov exponent (LLE),[18,30,38,39] entropy-based,[11,39] coherence,[17,27,29,31] synchronization,[32–34,36] fractal dimension,[11,26,35] frequency,[13] higher order spectra,[11,22,39,40] power density spectrum,[15,16] relative convergence,[37] empirical mode decomposition (EMD),[21] and intrinsic time-scale decomposition (ITD).[23] In addition, statistical analysis, e.g. analysis of variation (ANOVA),[11,27,33,36,40] and some machine learning-based classifiers[11,15,16,18,21,22,28,29,39,41] are used for feature discrimination and pattern identification. Another well-known approach of EEG analysis is the wavelet-chaos methodology.[14,18,26,30,35,38] Based on a review of the related literatures, until now, there is no effective method to predict the refractoriness in patients with idiopathic seizures, especially based on EEG analysis.

In this study, we attempted to develop an efficient, automated, and quantitative approach for early prediction of refractory idiopathic epilepsy based on well-controlled/refractory EEG classification analysis. In the proposed approach, a set of artifact-free EEG segments was acquired from the first EEG recordings of two classes of patients,

namely, well-controlled and refractory. Note that all EEG recordings were performed before starting AED treatments on the patients. To search for the discriminative and significant EEG features and reduce the computation cost for the classification analysis, an approach involving global parametric features (GPF)[12] was adopted for statistically characterizing EEG segments across channels as well as over time, and a gain ratio-based feature selection was then performed. The advantage of using GPF was its capability of statistically characterizing and comparing EEG segments regardless of the different recording durations and number of channels. In addition, the gain ratio-based feature selection was adopted because of its low computation complexity. After that, a well-known classification model based on the support vector machine (SVM) was trained with the corresponding selected feature data from EEG segments.[42–44] The trained SVM was finally employed to classify an unknown patient as well-controlled or refractory. The utilization of SVM was due to its excellent discrimination and generalization capabilities and processing efficiency for high dimensional data. To increase the efficiency and reliability, some functionalities of two well-known free software packages, EPILAB and Weka,[45,46] were integrated into the implementation of our approach. The former was applied for EEG feature extraction, while the latter was applied for EEG feature selection and SVM-based classification model construction. We attempted to predict refractory idiopathic epilepsy at an early stage in the current study.

2. Methods

2.1. *Subjects*

To minimize the impact of EEG factors on refractory epilepsy prediction, 23 Taiwanese children (nine boys and 14 girls) who had epileptiform discharges in their EEG examinations and were diagnosed with idiopathic epilepsy were enrolled, as shown in Table 1. The diagnosis of epilepsy was made according to the criteria established by the International League Against Epilepsy (ILAE). Idiopathic epilepsy is defined as an epilepsy of predominately genetic or presumed genetic origin and in which there is no gross neuroanatomic or neuropathologic abnormality.[47] They were divided into two epilepsy groups, well-controlled and refractory. The numbers of patients in the well-controlled and refractory classes were 12 and 11, respectively. Participants were classified as having well-controlled epilepsy if they had achieved adequate seizure control relative to their epilepsy condition following the use of appropriate AED treatment.[48] In this study, no seizure was found for at least one year in well-controlled patients. Refractory epilepsy was defined as failure of seizure control, with more than two AED at maximum tolerated doses, with an average of more than one seizure per month for 18 months, and no more than three consecutive months seizure-free during this interval.[49] The mean age of seizure onset in the well-controlled group was 8 years ±2 years and in the refractory group was 9 years 9 months ±4 years. A written informed consent was given by a family member or legal guardian in each case. This study was approved by the Institutional Review Board of the Kaohsiung Medical University Hospital (KMUIRB-2012-01-09 (II), KMUHIRB-20120095).

2.2. *Electroencephalogram examinations*

The patient was placed in a dark and quiet room and scalp EEG was recorded over a period of time, usually 20 min. The first EEG of each patient was analyzed before starting an AED treatment when they first visited our pediatric neurological department to exclude a drug effect. To minimize the postictal effect, patients received EEG examinations at least five days after a seizure attack. Each EEG was recorded digitally using 21 electrodes with a 200 Hz sampling rate (Harmonie DVN V5.1, Montreal, Canada). Amplifier characteristics were: 10,000 times gain, low-pass filter at 60 Hz, and high-pass filter at 0.5 Hz. Electrodes were placed according to the International 10-20 System. All recordings were performed during the daytime. To decrease the factors influencing EEG data, each patient maintained the same state of wakefulness throughout the recording period. Only EEG segments in which the patient was awake were analyzed (presence of alpha activity, no segments with decreased alpha activity, increased theta activity, nor vertex sharp waves).[50] For an unbiased comparison of EEGs, patients who had a seizure attack during the recording period were excluded.[12]

Table 1. Clinical characteristics of patients.

Patient number	Sex	Present AED	Age of first EEG	Seizure type	CT or MRI	Seizure before first EEG
			Well-controlled			
1	M	OXC	10y	Focal	CT:NP	2
2	M	nil	8y4m	Focal	CT:NP	1
3	F	nil	7y1m	Generalized	ND	1
4	F	nil	10y4m	Focal	ND	2
5	M	nil	6y9m	Generalized	MRI:NP	1
6	F	nil	5y9m	Focal	ND	1
7	M	OXC	7y8m	Focal	CT:NP	3
8	F	OXC	8y9m	Generalized	ND	1
9	F	VPA	11y6m	Generalized	ND	3
10	F	CBZ	6y8m	Focal	CT:NP	1
11	F	OXC	4y7m	Focal	MRI:NP	3
12	M	OXC	8y11m	Focal	ND	2
			Refractory			
13	F	LTG + LEV	9y	Focal	MRI:NP	2
14	M	CBZ + OXC + TPX	4y9m	Focal	MRI:NP	2
15	F	VPA + OXC + LEV + NIT	7y8m	Focal	MRI:NP	4
16	F	OXC + TPX + LEV + CLO	9y	Generalized	CT:NP, MRI:NP	3
17	M	CBZ + VPA + CLO1 + LEV	16y3m	Focal	CT:NP	1
18	F	OXC + VIG + CLO	13y8m	Focal	CT:NP	5
19	F	VPA + LTG	13y2m	Generalized	MRI:NP	2
20	F	VPA + OXC	13y	Generalized	ND	2
21	F	OXC + LEV	5y11m	Focal	MRI:NP	3
22	M	VPA + LEV + OXC	10y11m	Generalized	MRI:NP	1
23	M	CBZ + TPX + CLO	4y2m	Generalized	CT:NP	1

Note: M: male, F: female, nil: No AED use, ND: not done, NP: normal, CBZ: carbamazepine, OXC: oxcarbazepine, VPA: valproic acid, LTG: lamotrigine, VIG: vigabatrin, NIT: nitrazepam, TPM: topamax, CLO: clonazepam, LEV: levetiracetam.

Patients were not under any AED before their first EEG recordings.

2.3. *EEG classification analysis*

In this study, an approach was developed for well-controlled/refractory classification by following the standard procedure of pattern recognition,[51,52] as outlined in Fig. 1. Two phases were included in the approach, the training phase and the classification phase. The training phase was mainly composed of five steps, namely EEG acquisition, feature extraction, feature transformation, feature selection, and SVM training. First, two classes of artifact-free EEG segments were acquired from the first EEG examinations of well-controlled and refractory patients, respectively. Then, 24 existing EEG features were extracted from each EEG segment and further transformed into 216 global feature descriptors. Next, a gain ratio-based feature selection[51] was performed to select the best 10 discriminative descriptors. Finally, the set of EEG segments characterized with the selected global feature descriptors and labeled with the corresponding classes was employed for training a SVM. After the training phase, the trained SVM was employed to classify an unknown patient as well-controlled or refractory in the classification phase. Below, each step is described in greater detail.

2.3.1. *EEG acquisition*

As mentioned above, 23 patients were enrolled in this study and divided into two classes, well-controlled and refractory, according to the ILAE criteria. A dataset, comprised of 16 well-controlled and 14 refractory segments, was acquired from the 23 enrolled patients. Note that for each patient, one, two, or three segments were acquired from his/her

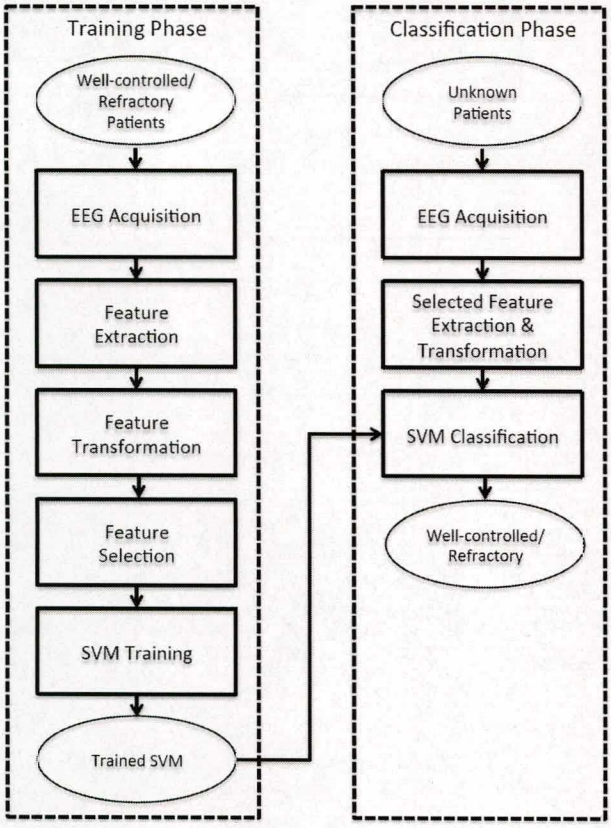

Fig. 1. Flowchart of our approach.

first EEG recording made before starting an AED treatment. A total of eight channels of longitudinal bipolar montage were adopted for the analysis, including F3-C3, F4-C4, C3-T3, C4-T4, T3-O1, T4-O2, O1-C3, and O2-C4. For an unbiased comparison, all EEG segments were acquired from artifact-free sections of awake recordings. In addition, the size of each segment ranged from 66 to 547 s. The dataset was denoted with $\{(E_i, C_i) \mid i = 1, 2, \ldots, 30\}$, where E_i was the $8 \times n_i$ matrix of the ith EEG segment with 8 bipolar channels and n_i samples and $C_i \in \{\textit{well-controlled}, \textit{refractory}\}$ was the corresponding class label of E_i.

2.3.2. *Feature extraction*

To characterize each EEG segment, a software package, i.e., EPILAB,[45] was applied to extract univariate (one-channel-based) EEG features from the segment in the feature extraction step. EPILAB is a free Matlab-based software package developed for EEG-based epileptic seizure prediction by the European project EPILEPSIAE. To search for the crucial EEG features as comprehensively as possible, a total of 24 univariate EEG features were considered and were divided into 9 categories, including AR modeling predictive error, decorrelation time, energy, entropy, Hjorth, relative power, spectral edge, statistics, and energy of the wavelet coefficients, as shown in Table 2. These features were computed for each bipolar channel of each segment in a window-by-window manner, as illustrated in Fig. 2. The window size was set to 5 s as suggested in EPILAB. Therefore, a dataset of EEG feature matrices with corresponding class labels was obtained, $F = \{(F_{i1}, F_{i2}, \ldots, F_{i24}, C_i) \mid i = 1, 2, \ldots, 30\}$, where F_{ij} was the corresponding $8 \times n'_i$ matrix of the jth feature of E_i with 8 bipolar channels and $n'_i = \lfloor n_i/(5 \times 200) \rfloor$ time windows and represented as the following equation:

$$F_{ij} = \begin{bmatrix} f_{ij}(1,1) & f_{ij}(1,2) & \cdots & f_{ij}(1,n'_i) \\ f_{ij}(2,1) & f_{ij}(2,2) & \cdots & f_{ij}(2,n'_i) \\ \vdots & \vdots & \ddots & \vdots \\ f_{ij}(8,1) & f_{ij}(8,2) & \cdots & f_{ij}(8,n'_i) \end{bmatrix}. \quad (1)$$

Note that $\lfloor n_i/(5 \times 200) \rfloor$ denoted the floor function which mapped $n_i/(5 \times 200)$ to the largest previous integer n'_i. In addition, $f_{ij}(l,k)$ denoted the feature

Table 2. EEG feature categories and the all corresponding features screened in current study.

Feature categories	Features
AR modelling predictive error	AR modelling predictive error
Decorrelation time	Decorrelation time
Energy	Energy
Entropy	Approximate entropy, Sample entropy
Hjorth	Mobility, Complexity
Relative power	Delta band (0.1–4 Hz), Theta band (4–8 Hz), Alpha band (8–15 Hz), Beta band (15–30 Hz), Gamma band (30–2000 Hz)
Spectral edge	Power, Frequency
Statistics	1st moment (mean), 2nd moment (variance), 3rd moment (skewness), 4th moment (kurtosis)
Energy of the wavelet coefficients	Energy of Daubechies order 4 wavelet transform in decomposition levels 1, 2, 3, 4, 5, 6

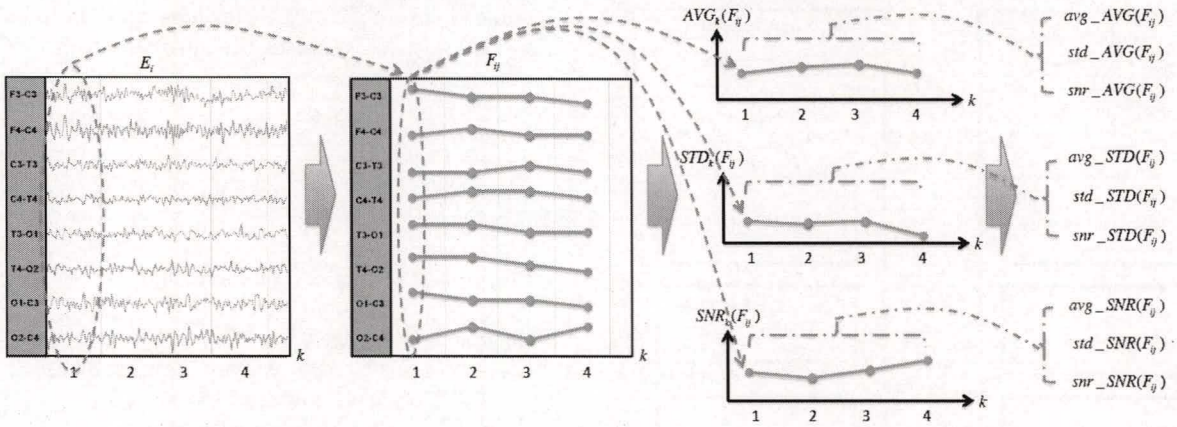

Fig. 2. Illustration of feature extraction and transformation.

value element of F_{ij} and corresponded to the kth time window in the lth bipolar channel.

2.3.3. *Feature transformation*

To describe the distribution of feature values across 8 bipolar channels as well as over n'_i time windows in each feature matrix F_{ij}, an approach involving GPF[12] was adopted to transform each feature matrix F_{ij} into a matrix of global feature descriptors in the feature transformation step. The advantage of GPF was its capability of statistically characterizing and comparing EEG segments regardless of the different recording durations and number of channels. As illustrated in Fig. 2, for each feature matrix F_{ij}, we first calculated the inter-channel average (AVG), inter-channel standard deviation (STD), and inter-channel signal to noise ratio (SNR) of feature values across eight bipolar channels in the kth time window with the following equations:

$$\mathrm{AVG}_k(F_{ij}) = \frac{1}{8}\sum_{l=1}^{8} f_{ij}(l,k), \tag{2}$$

$$\mathrm{STD}_k(F_{ij}) = \sqrt{\frac{1}{8}\sum_{l=1}^{8}(f_{ij}(l,k) - \mathrm{AVG}_k(F_{ij}))^2}, \tag{3}$$

$$\mathrm{SNR}_k(F_{ij}) = \frac{\mathrm{AVG}_k(F_{ij})}{\mathrm{STD}_k(F_{ij})}. \tag{4}$$

After that, we further calculated the average over time (avg), standard deviation over time (std), and signal to noise ratio over time (snr) of AVGs, STDs, and SNRs over n'_i time windows with the following equations:

$$\mathrm{avg\_AVG}(F_{ij}) = \frac{1}{n'_i}\sum_{k=1}^{n'_i} \mathrm{AVG}_k(F_{ij}), \tag{5}$$

$$\mathrm{avg\_STD}(F_{ij}) = \frac{1}{n'_i}\sum_{k=1}^{n'_i} \mathrm{STD}_k(F_{ij}), \tag{6}$$

$$\mathrm{avg\_SNR}(F_{ij}) = \frac{1}{n'_i}\sum_{k=1}^{n'_i} \mathrm{SNR}_k(F_{ij}), \tag{7}$$

$$\mathrm{std\_AVG}(F_{ij}) = \sqrt{\frac{1}{n'_i}\sum_{k=1}^{n'_i}(\mathrm{AVG}_k(F_{ij}) - \mathrm{avg\_AVG}(F_{ij}))^2}, \tag{8}$$

$$\mathrm{std\_STD}(F_{ij}) = \sqrt{\frac{1}{n'_i}\sum_{k=1}^{n'_i}(\mathrm{STD}_k(F_{ij}) - \mathrm{avg\_STD}(F_{ij}))^2}, \tag{9}$$

$$\mathrm{std\_SNR}(F_{ij}) = \sqrt{\frac{1}{n'_i}\sum_{k=1}^{n'_i}(\mathrm{SNR}_k(F_{ij}) - \mathrm{avg\_SNR}(F_{ij}))^2}, \tag{10}$$

$$\mathrm{snr\_AVG}(F_{ij}) = \frac{\mathrm{avg\_AVG}(F_{ij})}{\mathrm{std\_AVG}(F_{ij})}, \tag{11}$$

$$\mathrm{snr\_STD}(F_{ij}) = \frac{\mathrm{avg\_STD}(F_{ij})}{\mathrm{std\_STD}(F_{ij})}, \tag{12}$$

$$\mathrm{snr\_SNR}(F_{ij}) = \frac{\mathrm{avg\_SNR}(F_{ij})}{\mathrm{std\_SNR}(F_{ij})}. \tag{13}$$

Thus, each feature matrix F_{ij} was transformed into a global feature descriptor matrix GF_{ij} as

$$\mathrm{GF}_{ij} = \begin{bmatrix} \text{avg\_AVG}(F_{ij}) & \text{std\_AVG}(F_{ij}) & \text{snr\_AVG}(F_{ij}) \\ \text{avg\_STD}(F_{ij}) & \text{std\_STD}(F_{ij}) & \text{snr\_STD}(F_{ij}) \\ \text{avg\_SNR}(F_{ij}) & \text{std\_SNR}(F_{ij}) & \text{snr\_SNR}(F_{ij}) \end{bmatrix}. \tag{14}$$

Finally, a dataset of transformed EEG global feature descriptor matrices with the corresponding class labels was obtained $\{(\mathrm{GF}_{i1}, \mathrm{GF}_{i2}, \ldots, \mathrm{GF}_{i24}, C_i) \mid i = 1, 2, \ldots, 30\}$.

2.3.4. *Feature selection*

In the previous step, we had obtained 24 transformed matrices, each of which contained nine global feature descriptors for each EEG segment, resulting in a total of 216 global feature descriptors. These descriptors were denoted as $\mathrm{gfd}_1, \mathrm{gfd}_2, \ldots, \mathrm{gfd}_{216}$. Among these descriptors, some may be redundant and/or may not contain enough discriminative information for the well-controlled/refractory classification. In addition, a classifier with a high number of features may suffer from the curse of dimensionality. Therefore, we applied the well-known free data mining software, Weka,[46] for selecting discriminative ones from the 216 global feature descriptors. There are several feature selection measures developed in Weka. According to our experiments, the gain ratio measure was adopted due to the low computation cost and satisfactory performance evaluation.[51] For each global feature descriptor gfd_s, the corresponding gain ratio GainRatio (gfd_s) was calculated using the following equation:

$$\text{GainRatio } (\mathrm{gfd}_s) = \frac{\text{Gain } (\mathrm{gfd}_s)}{\text{SplitInfo } (\mathrm{gfd}_s)}, \tag{15}$$

where Gain (gfd_s) denoted the information gain which was a measure of the change in information entropy after splitting the dataset into subsets with the global feature descriptor gfd_s; SplitInfo (gfd_s) denoted the split information which was a measure of the information entropy generated by splitting the dataset into subsets with gfd_s, as defined in the following equation:

$$\text{SplitInfo } (\mathrm{gfd}_s) = -\sum_{j=1}^{p} \frac{|D_j|}{|D|} \times \log_2 \left(\frac{|D_j|}{|D|} \right), \tag{16}$$

where D denotes the dataset, $D_1, D_2, \ldots, D_p$ denotes the mutually disjoint partitions of D, p denotes the number of partitions of D, $|D|$ and $|D_j|$ denote the corresponding number of data included in D and D_j respectively. Note that information entropy is a measure of the uncertainty or unpredictability of information content in the information theory. It is usually applied to measure the "impurity" of a dataset with respect to classes in classification analysis. The larger the gain ratio was, the more discriminative the gfd_s was. More details for the overview of gain ratio is described by Han *et al.*[52] According to the ranking for each descriptor provided by the gain ratio measure, the descriptors having the best 10 scores for the measure were chosen as the selected ones for each EEG segment. Note that more descriptors selected usually increase the computation cost while less descriptors selected may decrease the discriminative capability. Based on the results of our experiments, selecting the best 10 descriptors was efficient and acceptable for the well-controlled/refractory classification. Table 3 presents the best 10 descriptors and the corresponding gain ratios. Three descriptors, including DecorrTime_avg_AVG, DecorrTime_std_AVG, and DecorrTime_snr_STD, were related to "decorrelation time," defined as the time of the first zero crossing of the autocorrelation sequence of the EEG signal in a considered window; four descriptors, RelPowDelta_avg_AVG, RelPowDelta_avg_SNR, RelPowDelta_snr_AVG, and RelPowGamma_std_SNR, were related to "relative power of delta/gamma," defined as the relative power of delta/gamma band in all spectral bands, including delta, theta, alpha, beta, and gamma, of the EEG signal in a considered window; one descriptor, SpectrEdgeFreq_avg_AVG, was related to "spectral edge frequency," defined as the minimal frequency below which 50% of the total power of the EEG signal in a considered window were located; one descriptor, Kurtosis_snr_STD, was related to "kurtosis," defined as a quantification of the relative peakness of the amplitude distribution of the EEG signal in a considered window; one descriptor, Wavelet_db4_EnergyBand_5_snr_STD, was defined as the energy of Daubechies order 4 wavelet transform in decomposition level 5. Note that the string concatenating in the rear of each feature name, namely "str_STR," where str $\in \{\text{avg}, \text{std}, \text{snr}\}$ and STR $\in \{\text{AVG}, \text{STD}, \text{SNR}\}$, denoted the average, standard deviation, or signal to noise ratio over time of the inter-channel averages, standard deviations, or signal

Table 3. Selected global feature descriptors and the corresponding ranking based on gain ratio measure.

Ranking	Features (abbr.)	Features (full)	Gain ratio
1	DecorrTime_avg_AVG	Decorrelation Time_avg_AVG	0.538
2	RelPowDelta_avg_AVG	Relative Power of Delta_avg_AVG	0.538
3	DecorrTime_std_AVG	Decorrelation Time_std_AVG	0.450
4	SpectrEdgeFreq_avg_AVG	Spectral Edge Frequency_avg_AVG	0.432
5	DecorrTime_snr_STD	Decorrelation Time_snr_STD	0.405
6	RelPowDelta_avg_SNR	Relative Power of Delta_avg_SNR	0.405
7	RelPowGamma_std_SNR	Relative Power of Gamma_std_SNR	0.405
8	Kurtosis_snr_STD	Kurtosis_snr_STD	0.363
9	Wavelet_db4_EnergyBand_5_snr_STD	Wavelet_db4_EnergyBand_5_snr_STD	0.327
10	RelPowDelta_snr_AVG	Relative Power of Delta_snr_AVG	0.326

to noise ratios of the feature. By letting sf_{ij} denote the value of jth selected global feature descriptor for the ith EEG segment, a dataset of selected global feature descriptor vectors with the corresponding class labels was obtained $\{(sf_{i1}, sf_{i2}, \ldots, sf_{i10}, C_i)\ i = 1, 2, \ldots, 30\}$.

2.3.5. *SVM training*

To build a model for well-controlled/refractory classification, ν-SVM,[42] a well-known algorithm of SVMs, was adopted. SVMs are supervised learning models with associated learning algorithms for pattern classification and regression analysis. For a two-class classification problem, the basic idea of SVMs is to project the training data with kernels into a higher dimension space where the data are linearly separable and then to find an optimized hyperplane that separates the two classes of data and maximizes the distance to the nearest data point, the so-called margin. To resolve the restricted constraint of no error on the training data in traditional SVMs and gain the control of generalization capability for test data, a parameter ν was added into the ν-SVM. Therefore, ν-SVM was more suitable for complex classification problems.[53] In this study, we employed the LibSVM,[54] which is a library for support vector machines and has been integrated into the Weka, to run the ν-SVM algorithm on the selected feature dataset $\{(sf_{i1}, sf_{i2}, \ldots, sf_{i10}, C_i) \,|\, i = 1, 2, \ldots, 30\}$. The kernel type was chosen as radial basis function, and the related parameters of ν-SVM, including nu (ν), gamma (g), cost (c), and epsilon (p) were set as 0.15, 0.1, 1.0, and 0.001, respectively. In addition, the 10-fold cross-validation mode was adopted for assessing how the trained ν-SVM model would generalize to a nontraining dataset. The n-fold cross-validation is one of the most popular techniques for estimating the prediction error of a classifier and has also been widely utilized in many studies of medical analysis and applications.[55–58] As suggested by Kohavi[59] and Rodriguez *et al.*[60] the value of n is optimally set as 10, yielding a 10-fold cross-validation. The dataset was randomly partitioned into 10 equal size subsets, each of which contained three data points. Of the 10 subsets, one was retained as the validation data for testing the model, and the remaining nine subsets were used as training data. The process was repeated 10 times with each of the 10 subsets retained exactly once as the validation data, yielding a total of 30 validation data. To evaluate the accuracy of the classification model and results, several performance indices supported by Weka were calculated, including TP (true positive) rate, FP (false positive) rate, precision, recall, F-measure, MCC (Matthews correlation coefficient),[61] and ROC (Receiver operating characteristics) area.[62] Let the considered class, well-controlled or refractory, be positive and the other one negative. A TP/FP occurred when the data was correctly/incorrectly classified into the positive class and it actually belonged to positive/negative class, while a TN (true negative)/FN (false negative) occurred when the data was correctly/incorrectly classified into the negative class and it actually belonged to negative/positive class. Accordingly, TP rate, FP rate, precision, recall, and F-measure were defined by the following equations:

$$\text{TP rate} = \frac{\text{num(TP)}}{\text{num(TP)} + \text{num(FN)}}, \tag{17}$$

$$\text{FP rate} = \frac{\text{num(FP)}}{\text{num(FP)} + \text{num(TN)}}, \tag{18}$$

$$\text{precision} = \frac{\text{num(TP)}}{\text{num(TP)} + \text{num(FP)}}, \tag{19}$$

$$\text{recall} = \frac{\text{num(TP)}}{\text{num(TP)} + \text{num(FN)}}, \tag{20}$$

$$F = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}, \tag{21}$$

where num(·) was the number of the corresponding event. MCC was a measure of the quality of two-class classification and was defined as the following equations[61]:

$$\text{MCC} = \frac{\text{num(TP)} \times \text{num(TN)} - \text{num(FP)} \times \text{num(FN)}}{\sqrt{B_1 \times B_2}}, \tag{22}$$

$$B_1 = (\text{num(TP)} + \text{num(FP)}) \times (\text{num(TP)} + \text{num(FN)}), \tag{23}$$

$$B_2 = (\text{num(TN)} + \text{num(FP)}) \times (\text{num(TN)} + \text{num(FN)}). \tag{24}$$

ROC area was a measure of the area under a ROC curve. It ranged between 0 and 1 and represented the probability that the classifier will rank a randomly chosen positive instance higher than a randomly chosen negative instance. More details about ROC area is described by Fawcett.[62] According to the definitions of measures mentioned above, a better classification model should produce a lower FP rate and higher TP rate, precision, recall, F-measure, MCC, and ROC area.

2.4. *Statistical analysis*

Data are shown as means ± SD. Distributions of seizure type, age of first EEG, seizure number before diagnosis, and gender between well-controlled and refractory groups were compared by Chi-square tests. The Wilcoxon rank sum test was performed to verify the statistical significance of the selected 10 features between the well-controlled and refractory groups. A p-value of less than 0.05 was considered statistically significant. All statistical analyses were conducted by SAS JMP (SAS Institute Inc., Cary, US).

3. Results

As shown in Table 1, 12 patients with well-controlled epilepsy and 11 patients with refractory epilepsy were enrolled. Four of the 12 well-controlled patients were classified as having generalized seizures (33.3%) and the others had focal seizures; five of the 11 refractory patients had generalized seizures (45.4%) and the others had focal seizures. All of the patients were idiopathic in etiology and had normal IQ. In the well-controlled group, seven patients received one kind of AED and five patients did not receive AED. In the refractory group, all of the patients received more than two AEDs. Brain computed tomography (CT) or Magnetic resonance imaging (MRI) were performed in 10 of the 11 refractory patients, and all of them showed unremarkable results. Six of the 12 well-controlled patients received brain CT or MRI examinations, and revealed no definite intracranial lesion. Before their first EEG examinations, the average number of seizure episodes were 1.8 ± 0.9 and 2.4 ± 1.3 times in the well-controlled and refractory groups, respectively. There were no significant differences in clinical characteristics, including age of the first EEG, gender, seizure type, and seizure number before diagnosis, between the well-controlled and refractory groups.

3.1. *Selected features with corresponding ranking based on gain ratio*

There were 10 crucial feature descriptors selected for the well-controlled/refractory classification, including

DecorrTime_avg_AVG, RelPowDelta_avg_AVG, DecorrTime_std_AVG, SpectrEdgeFreq_avg_AVG, DecorrTime_snr_STD, RelPowDelta_avg_SNR, RelPowGamma_std_SNR, Kurtosis_snr_STD, Wavelet_db4_EnergyBand_5_snr_STD, and RelPowDelta_snr_AVG,

which were ranked based on gain ratio as shown in Table 3. Gain ratio values ranged from 0.326 to 0.538. As mentioned earlier, the larger the gain ratio was, the more discriminative for well-controlled/refractory classes the global feature descriptor was. Figure 3 shows the data distribution of 30 EEG segments acquired from 12 well-controlled patients, $\{Pi \,|\, 1 \leq i \leq 12\}$, and 11 refractory patients, $\{Pi \,|\, 13 \leq i \leq 23\}$ on the feature space generated by the best two descriptors, i.e., DecorrTime_avg_AVG and RelPowDelta_avg_AVG. Note that the string "$-j$" concatenating in the rear of Pi denotes the jth EEG segment acquired from patient i. In

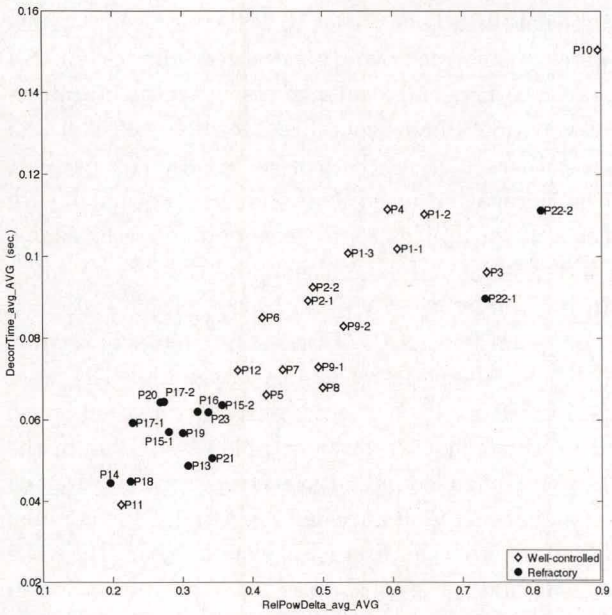

Fig. 3. Distribution of "RelPowDelta_avg_AVG" and DecorrTime_avg_AVG" feature values calculated from all 30 EEG segments of 23 patients. Note that the string "-j" concatenating in the rear of Pi denotes the jth EEG segment acquired from patient i. Besides, the data labeled by "hollow diamond" and "solid circle" correspond to well-controlled and refractory patients, respectively.

addition, the data labeled by "hollow diamond" and "solid circle" correspond to well-controlled and refractory patients, respectively. Apparently, the majority of data from the well-controlled patients demonstrated higher DecorrTime_avg_AVG and RelPowDelta_avg_AVG, while the majority of data from refractory patients showed lower DecorrTime_avg_AVG and RelPowDelta_avg_AVG. In other words, the set of the best two selected global feature descriptors truly possessed good discriminability for well-controlled/refractory classification. However, a few data points were far away from the corresponding groups of data, namely P11, P22-1, and P22-2. Therefore, more additional global feature descriptors were required to increase discriminability.

3.2. *Features comparison between well-controlled and refractory groups*

To verify the statistical significance of the selected 10 global feature descriptors, the Wilcoxon rank sum test was used to compare these descriptors between well-controlled and refractory groups. There were significantly higher DecorrTime_avg_AVG (0.088 ± 0.025 versus $0.063 \pm 0.018, p = 0.0015$), RelPowDelta_avg_AVG ($0.523 \pm 0.155$ versus $0.356 \pm 0.183, p = 0.0030$), DecorrTime_std_AVG ($0.019 \pm 0.006$ versus $0.013 \pm 0.008, p = 0.0261$), DecorrTime_snr_STD ($2.753 \pm 1.052$ versus $2.278 \pm 1.618, p = 0.0358$), RelPowDelta_avg_SNR ($6.581 \pm 6.194$ versus $4.634 \pm 5.239, p = 0.0065$), RelPowGamma_std_SNR ($0.656 \pm 0.252$ versus $0.491 \pm 0.244, p = 0.0483$), and RelPowDelta_snr_AVG (8.590 ± 8.433 versus 4.640 ± 1.589, $p = 0.0065$) in the well-controlled group than in the refractory group. On the contrary, there were significantly lower SpectrEdgeFreq_avg_AVG (4.232 ± 1.145 versus 5.530 ± 1.338, $p = 0.0083$), Kurtosis_snr_STD (1.298 ± 0.327 versus $1.694 \pm 0.306, p = 0.0044$), and Wavelet_db4_EnergyBand_5_snr_STD (1.900 ± 0.788 versus 2.497 ± 0.617, $p = 0.0261$) in the well-controlled group than in the refractory group. Since all p-values in these tests were less than 0.05, the selected 10 global feature descriptors were considered statistically significant. We selected the best two feature descriptors to represent the difference between the well-controlled and refractory groups. Figure 4 demonstrates a comparison of "DecorrTime" and "RelPowDelta" feature values in all bipolar channels between a well-controlled patient (P1-2) and a refractory patient (P14). Also, the corresponding raw EEGs in a selected time interval [60, 70] from these two patients' EEG segments are also presented. From visual inspection of raw EEGs (Fig. 4(a)), it may be difficult to distinguish the well-controlled patient from the refractory patient. In addition, it is clear that there are more frequent epileptiform discharges are found in the raw EEG of the well-controlled patient than that of the refractory patient. However, by QEEG analysis (Fig. 4(b)), the values and variations of Decorrelation Time and Relative Power of Delta in all bipolar channels over time are significantly higher in the well-controlled patient (P1-2) than those in the refractory patient (P14). This observation conforms with the statistical results mentioned above.

3.3. *The classification results*

As mentioned earlier, a 10-fold cross-validation mode, adopted for assessing how the trained ν-SVM model would generalize to a nontraining dataset, yielded a total of 30 validation data. The set of validation data contained 16 data of EEG segments

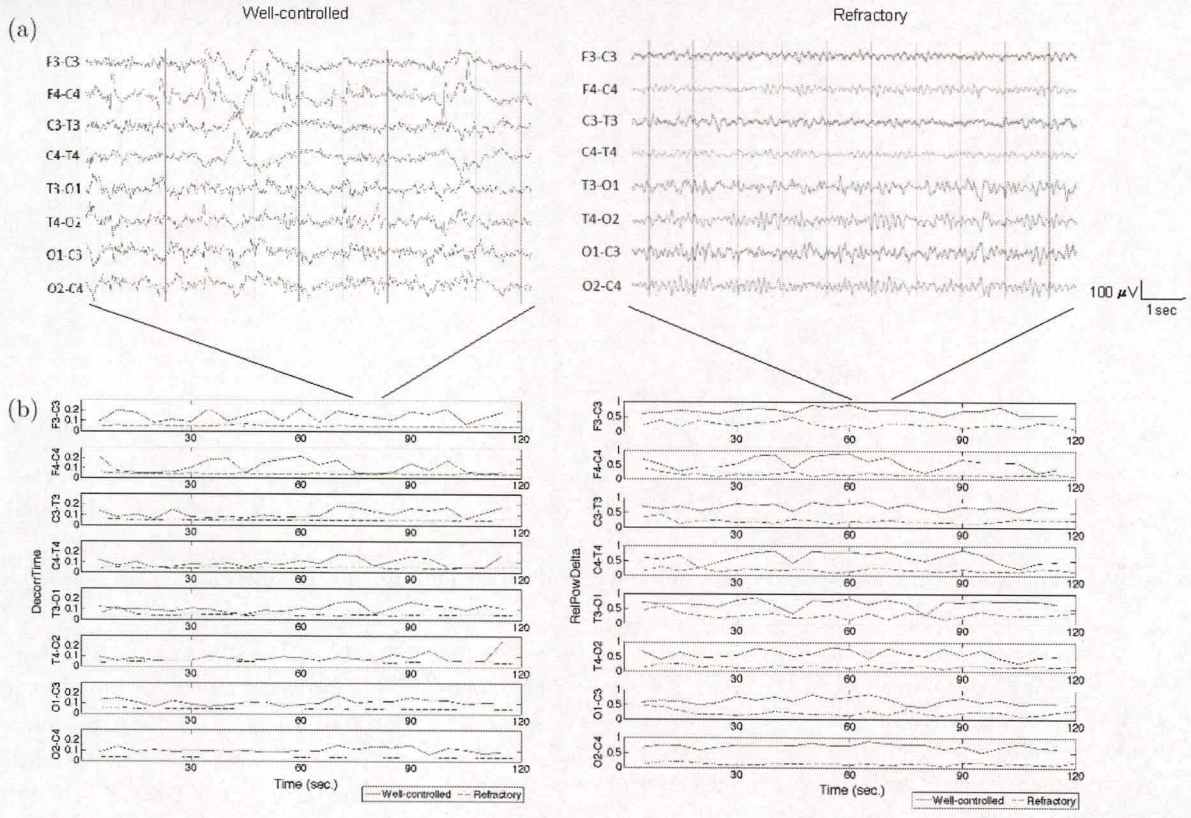

Fig. 4. A comparison of "DecorrTime" and "RelPowDelta" feature values in all bipolar channels and the corresponding selected raw EEGs between a well-controlled patient (P1-2) and a refractory patient (P14).

from 12 well-controlled patients and 14 data of EEG segments from 11 refractory patients. For the 16 well-controlled EEG segments, 14 segments were correctly classified into the desired well-controlled class, while only two segments, P2-2 from patient 2 and P11 from patient 11, were incorrectly classified into the refractory class. However, there was no observed difference clinically between these two patients and the other patients in the well-controlled group, except that patient 11 was the youngest. Furthermore, all the refractory segments were correctly classified into the desired refractory class. Several performance indices for evaluating the classification result are listed in Table 4. Note that the index values listed in the second and third row are calculated for viewing the well-controlled class and refractory class as the positive one, respectively. In addition, each index value listed in the fourth row is the weighted average of the corresponding well-controlled and refractory values in the same column with weights 16 and 14, respectively. The results yield high weighted averages of TP rate (0.933), precision (0.942), recall (0.933), F-measure (0.933), MCC (0.875), ROC area (0.938), and a low weighted average of FP rate (0.058). In addition, 5-fold, 4-fold, 2-fold cross-validation modes were also performed and the corresponding number of incorrectly classified segments were two, four, and five, respectively, yielding the corresponding

Table 4. Performance indices for evaluating the classification result.

Class	TP rate	FP rate	Precision	Recall	F-measure	MCC	ROC area
Well-controlled	0.875	0.000	1.000	0.875	0.933	0.875	0.938
Refractory	1.000	0.125	0.875	1.000	0.933	0.875	0.938
Weighted Average	0.933	0.058	0.942	0.933	0.933	0.875	0.938

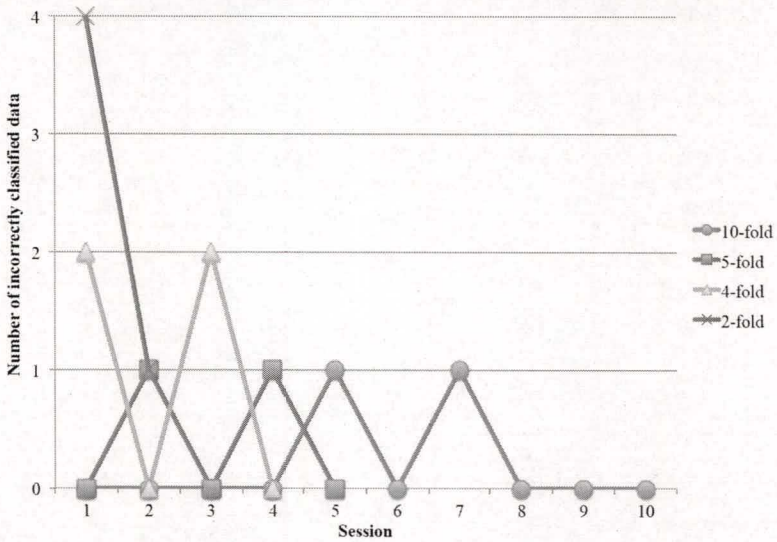

Fig. 5. A comparison on the inter-session variability of classification performance between four validation modes.

weighted averages of precision (0.933, 0.872, and 0.836, respectively), and the weighted averages of recall (0.933, 0.867, and 0.833, respectively). Figure 5 compares the inter-session variability of classification performance between the four validation modes. Compared with the validation results of 4-fold and 2-fold, the results of 10-fold and 5-fold presented less inter-session variability. The above results coincided with those suggested by Kohavi[59] and Rodriguez *et al.*[60] Apparently, the trained ν-SVM model possessed an excellent generalization capability on the validation data and thus can be employed for classifying an unknown patient as well-controlled or refractory efficiently and reliably.

4. Discussion

Early prediction of children at risk of developing refractory epilepsy is important for considering alternative treatments. Many studies have helped to identify a number of risk factors, such as remote symptomatic etiology, abnormal neuroimaging, a multiple first seizure, previous history of status epilepticus, an abnormal EEG and more than five seizures before diagnosis as significant predictors for developing refractory epilepsy.[8,63–65] Conversely, idiopathic etiology is considered a significant predictor of lower risk.[8] However, a number of patients with idiopathic epilepsy might be refractory to medical treatment.[9] In the current study, we successfully used a new EEG analytical method to predict refractory epilepsy at an early stage in most patients with idiopathic epilepsy. In addition, the 10 significant features we obtained are associated with seizure prediction or classification and have been reported in the literatures, including decorrelation time and spectral edge frequency for seizure prediction, and relative power for classification of refractory epilepsy.[66,67] Mormann *et al.* report that prior to seizures, there is a decrease in the power related to the lower frequencies of the EEG, which leads to a drop in the decorrelation time.[66] The deficit in absolute delta power is also found in a significant portion of refractory patients with epilepsy in a neurofeedback study.[67] In our study, there was significantly higher DecorrTime_avg_AVG and RelPowDelta_avg_AVG in the well-controlled group than in the refractory group. It is suggested that refractory patients have a higher risk of seizure attacks than well-controlled patients. The main contributions of this study are as follows: (1) the first development of an efficient, automated and quantitative approach and tool for early prediction of refractory idiopathic epilepsy based on EEG classification analysis; (2) identification of significant and important EEG features for discriminating between well-controlled idiopathic epilepsy and refractory idiopathic epilepsy.

QEEG analysis is a useful tool to evaluate background EEG activity and to facilitate subsequent

expert visual EEG interpretation.[68] In clinical settings, it is sufficiently clinically useful in spike and seizure detection, and spike dipole analysis.[68] There are many studies published regarding QEEG analysis that provide additional information for patients with epilepsy.[69–71] Santiago-Rodriguez *et al.* analyze the background EEGs from 18 patients with juvenile myoclonic epilepsy (JME). They report that patients with JME have normal background EEGs by visual inspection but have an increase in absolute power of delta, alpha and beta bands which are more evident in frontoparietal regions by QEEG analysis.[50] Although the neurophysiological meaning of an increase in absolute power has not yet been elucidated, it is postulated that the absolute power of bands in EEGs are associated with enhanced synchronization of different neuronal populations.[72,73] Neuronal network hypersynchronization is a fundamental mechanism in idiopathic generalized epilepsy.[74] Sackellares *et al.* also use a novel QEEG analysis and seizure detection algorithm to assist Intensive Care Unit staff in timely identification of nonconvulsive seizures.[75] EPILAB was developed for researchers in performing studies on the prediction of epileptic seizures based on QEEG and classification methods.[45] In this study, via data mining software, Weka, we built a discriminative method through feature transformation by integrating the statistical analysis in inter-channel and time series EEG. Then, refractoriness of seizure was classified based on the ν-SVM with weighted averages of precision rate and recall rate as high as 94.2% and 93.3% respectively. Moreover, weighted averages of FP rate (0.058) and the other performance indices, such as TP rate (0.933), F-measure (0.933), MCC (0.875), and ROC area (0.938) almost reached the optimal values 0 and 1 respectively. The good performance was mainly due to further feature transformation with GPF for statistically characterizing well-controlled/refractory EEG segments across channels as well as over time, highly significant selected global feature descriptors, and the ν-SVM with excellent discrimination and generalization capabilities. In addition, this study also revealed that the proposed approach possessed an efficient and reliable capability for early prediction of refractory idiopathic epilepsy.

EEG abnormalities, abnormal neuroimaging, and seizure frequencies before diagnosis are risk factors for developing refractory epilepsy.[8,63] In this study, all of the patients showed epileptiform discharges in their EEG examinations and all of the patients who received neuroimaging studies revealed unremarkable results. There was no significant difference in seizure number between the well-controlled and refractory groups. Consequently, the impact of the related risk factors, including EEG abnormalities, abnormal neuroimaging, and seizure number before first EEG, on our study was reduced.

Although the results demonstrated that the proposed method can provide a precise identification of refractoriness in patients with idiopathic seizures, this study at the present stage contains at least two limitations. First, the number of participants was somewhat limited. However, the statistical results are significant when the 10 global feature descriptors were compared between the well-controlled and refractory groups. Second, we only used pediatric subjects for analysis. Therefore, replications with more diverse subjects including pediatric and adult subjects in further studies may increase the reliability and generalization of our proposed method.

5. Conclusion

We developed an efficient, automated and quantitative approach for identifying the possibility of developing refractory epilepsy in patients with idiopathic epilepsy before initiating treatment with medications. By utilizing our developed tool, it will help identify patients who require more intensive investigation and treatment at an early stage.

Acknowledgments

We thank the families who participated in this study, Dr. John Ebinger for English editing, and the help we received from the Statistical Analysis Laboratory, Department of Medical Research, Kaohsiung Medical University Hospital, Kaohsiung Medical University. This study was supported partly by a grant from the Kaohsiung Medical University Hospital (KMUH101-1M36 and KMUH102-2T09) and grants from Ministry of Science and Technology, Taiwan (NSC 101-2221-E-214-070 and NSC 102-2221-E-214-053).

References

1. K. C. Nickels, B. R. Grossardt and E. C. Wirrell, Epilepsy-related mortality is low in children: A 30-year population-based study in Olmsted County, MN, *Epilepsia* **53**(12) (2012) 2164–2171.
2. P. Kwan and M. J. Brodie, Early identification of refractory epilepsy, *N. Engl. J. Med.* **342**(5) (2000) 314–319.
3. O. Devinsky, Patients with refractory seizures, *N. Engl. J. Med.* **340**(20) (1999) 1565–1570.
4. K. Vonck, M. Sprengers, E. Carrette, I. Dauwe, M. Miatton, A. Meurs, L. Goossens, D. E. H. V, R. Achten, E. Thiery, R. Raedt, V. A. N. R. D and P. Boon, A decade of experience with deep brain stimulation for patients with refractory medial temporal lobe epilepsy, *Int. J. Neural Syst.* **23**(1) (2013) 1250034.
5. S. D. Lhatoo, J. K. Solomon, A. W. McEvoy, N. D. Kitchen, S. D. Shorvon and J. W. Sander, A prospective study of the requirement for and the provision of epilepsy surgery in the United Kingdom, *Epilepsia* **44**(5) (2003) 673–676.
6. H. S. Wang and K. L. Lin, Ketogenic diet: An early option for epilepsy treatment, instead of a last choice only, *Biomed. J.* **36**(1) (2013) 16–17.
7. T. S. Ko and G. L. Holmes, EEG and clinical predictors of medically intractable childhood epilepsy, *Clin. Neurophysiol.* **110**(7) (1999) 1245–1251.
8. J. Ramos-Lizana, P. Aguilera-Lopez, J. Aguirre-Rodriguez and E. Cassinello-Garcia, Early prediction of refractory epilepsy in childhood, *Seizure* **18**(6) (2009) 412–416.
9. A. Cukiert, J. A. Burattini, P. P. Mariani, C. M. Cukiert, M. Argentoni-Baldochi, C. Baise-Zung, C. R. Forster and V. A. Mello, Outcome after extended callosal section in patients with primary idiopathic generalized epilepsy, *Epilepsia* **50**(6) (2009) 1377–1380.
10. E. Kharazmi, M. Peltola, M. Fallah, T. Keranen and J. Peltola, Idiopathic generalized epilepsies: A follow-up study in a single-center, *Acta Neurol. Scand.* **122**(3) (2010) 196–201.
11. U. R. Acharya, S. V. Sree, A. P. Alvin, R. Yanti and J. S. Suri, Application of non-linear and wavelet based features for the automated identification of epileptic EEG signals, *Int. J. Neural Syst.* **22**(2) (2012) 1250002.
12. M. Cabrerizo, M. Ayala, M. Goryawala, P. Jayakar and M. Adjouadi, A new parametric feature descriptor for the classification of epileptic and control EEG records in pediatric population, *Int. J. Neural Syst.* **22**(2) (2012) 1250001.
13. A. V. Medvedev, A. M. Murro and K. J. Meador, Abnormal interictal gamma activity may manifest a seizure onset zone in temporal lobe epilepsy, *Int. J. Neural Syst.* **21**(2) (2011) 103–114.
14. H. Adeli, Z. Zhou and N. Dadmehr, Analysis of EEG records in an epileptic patient using wavelet transform, *J. Neurosci. Methods* **123**(1) (2003) 69–87.
15. U. R. Acharya, S. V. Sree and J. S. Suri, Automatic detection of epileptic EEG signals using higher order cumulant features, *Int. J. Neural Syst.* **21**(5) (2011) 403–414.
16. O. Faust, U. R. Acharya, L. C. Min and B. H. Sputh, Automatic identification of epileptic and background EEG signals using frequency domain parameters, *Int. J. Neural Syst.* **20**(2) (2010) 159–176.
17. D. Sherman, N. Zhang, S. Garg, N. V. Thakor, M. A. Mirski, M. A. White and M. J. Hinich, Detection of nonlinear interactions of EEG alpha waves in the brain by a new coherence measure and its application to epilepsy and anti-epileptic drug therapy, *Int. J. Neural Syst.* **21**(2) (2011) 115–126.
18. S. Ghosh-Dastidar, H. Adeli and N. Dadmehr, Mixed-band wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection, *IEEE Trans. Biomed. Eng.* **54**(9) (2007) 1545–1551.
19. S. Ghosh-Dastidar and H. Adeli, Spiking neural networks, *Int. J. Neural Syst.* **19**(4) (2009) 295–308.
20. N. Mammone, D. Labate, A. Lay-Ekuakille and F. C. Morabito, Analysis of absence seizure generation using EEG spatial-temporal regularity measures, *Int. J. Neural Syst.* **22**(6) (2012) 1250024.
21. R. J. Martis, U. R. Acharya, J. H. Tan, A. Petznick, R. Yanti, C. K. Chua, E. Y. Ng and L. Tong, Application of empirical mode decomposition (EMD) for automated detection of epilepsy using EEG signals, *Int. J. Neural Syst.* **22**(6) (2012) 1250027.
22. U. R. Acharya, R. Yanti, Z. J. Wei, M. Muthu Rama Krishnan, T. J. Hong, R. J. Martis and L. C. Min, Automated diagnosis of epilepsy using CWT, HOS and texture parameters, *Int. J. Neural Syst.* **23**(3) (2013) 1350009.
23. R. J. Martis, U. R. Acharya, J. H. Tan, A. Petznick, L. Tong, C. K. Chua and E. Y. Ng, Application of intrinsic time-scale decomposition (ITD) to EEG signals for automated seizure prediction, *Int. J. Neural Syst.* **23**(5) (2013) 1350023.
24. W. Yu, Z. Weidong, Y. Qi, L. I. Xueli, M. Qingfang, Z. Xiuhe and W. Jiwen, Comparison of ictal and interictal EEG signals using fractal features, *Int. J. Neural Syst.* **23**(6) (2013) 1–11.
25. H. S. U. Wei-Yen, Application of quantum-behaved particle swarm optimization to motor imagery EEG classification, *Int. J. Neural Syst.* **23**(6) (2013) 1–15.
26. M. M. Ahmadlou, H. H. Adeli and A. A. Adeli, Fractality and a wavelet-chaos-methodology for EEG-based diagnosis of Alzheimer disease, *Alzheimer Dis. Assoc. Disord.* **25**(1) (2010) 85–92.
27. Z. Sankari, H. Adeli and A. Adeli, Intrahemispheric, interhemispheric, and distal EEG coherence

in Alzheimer's disease, *Clin. Neurophysiol.* **122**(5) (2011) 897–906.
28. M. Ahmadlou, H. Adeli and A. Adeli, New diagnostic EEG markers of the Alzheimer's disease using visibility graph, *J. Neural Transm.* **117**(9) (2010) 1099–1109.
29. Z. Sankari and H. Adeli, Probabilistic neural networks for diagnosis of Alzheimer's disease using conventional and wavelet coherence, *J. Neurosci. Methods* **197**(1) (2011) 165–170.
30. H. Adeli, S. Ghosh-Dastidar and N. Dadmehr, A spatio-temporal wavelet-chaos methodology for EEG-based diagnosis of Alzheimer's disease, *Neurosci. Lett.* **444**(2) (2008) 190–194.
31. Z. Sankari, H. Adeli and A. Adeli, Wavelet coherence model for diagnosis of Alzheimer disease, *Clin. EEG Neurosci.* **43**(4) (2012) 268–278.
32. M. Ahmadlou and H. Adeli, Functional community analysis of brain: A new approach for EEG-based investigation of the brain pathology, *Neuroimage* **58**(2) (2011) 401–408.
33. M. Ahmadlou and H. Adeli, Fuzzy synchronization likelihood with application to attention-deficit/hyperactivity disorder, *Clin. EEG Neurosci.* **42**(1) (2011) 6–13.
34. M. Ahmadlou, H. Adeli and A. Adeli, Graph theoretical analysis of organization of functional brain networks in ADHD, *Clin. EEG Neurosci.* **43**(1) (2012) 5–13.
35. M. Ahmadlou, H. Adeli and A. Adeli, Fractality and a wavelet-chaos-neural network methodology for EEG-based diagnosis of autistic spectrum disorder, *J. Clin. Neurophysiol.* **27**(5) (2010) 328–333.
36. M. Ahmadlou, H. Adeli and A. Adeli, Fuzzy Synchronization Likelihood-wavelet methodology for diagnosis of autism spectrum disorder, *J. Neurosci. Methods* **211**(2) (2012) 203–209.
37. M. Ahmadlou, H. Adeli and A. Adeli, Spatiotemporal analysis of relative convergence of EEGs reveals differences between brain dynamics of depressive women and men, *Clin. EEG Neurosci.* **44**(3) (2013) 175–181.
38. H. Adeli, S. Ghosh-Dastidar and N. Dadmehr, A wavelet-chaos methodology for analysis of EEGs and EEG subbands to detect seizure and epilepsy, *IEEE Trans. Biomed. Eng.* **54**(2) (2007) 205–211.
39. U. R. Acharya, S. V. Sree, S. Chattopadhyay and J. S. Suri, Automated diagnosis of normal and alcoholic EEG signals, *Int. J. Neural Syst.* **22**(3) (2012) 1250011.
40. U. R. Acharya, E. C. Chua, K. C. Chua, L. C. Min and T. Tamura, Analysis and automatic identification of sleep stages using higher order spectra, *Int. J. Neural Syst.* **20**(6) (2010) 509–521.
41. U. R. Acharya, S. V. Sree, S. Chattopadhyay, W. Yu and P. C. Ang, Application of recurrence quantification analysis for the automated identification of epileptic EEG signals, *Int. J. Neural Syst.* **21**(3) (2011) 199–211.
42. B. Schölkopf, A. J. Smola, R. C. Williamson and P. L. Bartlett, New support vector algorithms, *Neural Comput.* **12**(5) (2000) 1207–1245.
43. H. Dai, H. Zhang, W. Wang and G. Xue, Structural reliability assessment by local approximation of limit state functions using adaptive Markov chain simulation and support vector regression, *Comput.-Aided Civ. Infrastruct. Eng.* **27**(9) (2012) 676–686.
44. D. Li, L. Xu, E. D. Goodman, Y. Xu and Y. Wu, Integrating a statistical background-foreground extraction algorithm and SVM classifier for pedestrian detection and tracking, *Integr. Comput.-Aided Eng.* **20**(3) (2013) 201–216.
45. C. A. Teixeira, B. Direito, H. Feldwisch-Drentrup, M. Valderrama, R. P. Costa, C. Alvarado-Rojas, S. Nikolopoulos, M. Le Van Quyen, J. Timmer, B. Schelter and A. Dourado, EPILAB: A software package for studies on the prediction of epileptic seizures, *J. Neurosci. Methods* **200**(2) (2011) 257–271.
46. M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann and I. H. Witten, The WEKA data mining software: An update, *ACM SIGKDD Explorations Newsletter* **11**(1) (2009) 10–18.
47. S. D. Shorvon, The etiologic classification of epilepsy, *Epilepsia* **52**(6) (2011) 1052–1057.
48. M. Gandy, L. Sharpe, K. N. Perry, L. Miller, Z. Thayer, J. Boserio and A. Mohamed, Rates of DSM-IV mood, anxiety disorders, and suicidality in Australian adult epilepsy outpatients: A comparison of well-controlled versus refractory epilepsy, *Epilepsy Behav.* **26**(1) (2013) 29–35.
49. A. T. Berg, S. Shinnar, S. R. Levy, F. M. Testa, S. Smith-Rapaport and B. Beckerman, Early development of intractable epilepsy in children: A prospective study, *Neurology* **56**(11) (2001) 1445–1452.
50. E. Santiago-Rodriguez, T. Harmony, L. Cardenas-Morales, A. Hernandez and A. Fernandez-Bouzas, Analysis of background EEG activity in patients with juvenile myoclonic epilepsy, *Seizure* **17**(5) (2008) 437–445.
51. I. H. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques* (Morgan Kaufmann, San Francisco, 2005).
52. J. Han, M. Kamber and J. Pei, *Data Mining: Concepts and Techniques* (Morgan Kaufmann, San Francisco, 2011).
53. P. H. Chen, C. J. Lin and B. Schölkopf, A tutorial on-support vector machines, *Appl. Stoch. Models Bus. Ind.* **21**(2) (2005) 111–136.
54. C. C. Chang and C. J. Lin, LIBSVM: A library for support vector machines, *ACM Trans. Intell. Syst. Technol.* **2**(3) (2011) 1–27.
55. A. Takemura, A. Shimizu and K. Hamamoto, Discrimination of breast tumors in ultrasonic images

using an ensemble classifier based on the AdaBoost algorithm with feature selection, *IEEE Trans. Med. Imag.* **29**(3) (2010) 598–609.
56. Z. Aydin, Y. Altunbasak and H. Erdogan, Bayesian models and algorithms for protein β-sheet prediction, *IEEE/ACM Trans. Comput. Biol. Bioinf.* **8**(2) (2011) 395–409.
57. S. Siuly and Y. Li, Improving the separability of motor imagery EEG signals using a cross correlation-based least square support vector machine for brain–computer interface, *IEEE Trans. Neural Syst. Rehabil. Eng.* **20**(4) (2012) 526–538.
58. B. Taati, J. Snoek, D. Aleman and A. Ghavamzadeh, Data mining in bone marrow transplant records to identify patients with high odds of survival, *IEEE J. Biomed. Health Inform.* **18**(1) (2014) 21–27.
59. R. Kohavi, A study of cross-validation and bootstrap for accuracy estimation and model selection, in *Proc. 14th Int. Joint Conf. Artificial Intelligence* (Morgan Kaufmann, Montreal, Quebec, Canada, 1995), pp. 1137–1143.
60. J. D. Rodriguez, A. Perez and J. A. Lozano, Sensitivity analysis of k-fold cross validation in prediction error estimation, *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(3) (2010) 569–575.
61. B. W. Matthews, Comparison of the predicted and observed secondary structure of T4 phage lysozyme, *Biochim. Biophys. Acta* **405**(2) (1975) 442–451.
62. T. Fawcett, An introduction to ROC analysis, *Pattern Recog. Lett.* **27**(8) (2006) 861–874.
63. E. C. Wirrell, Predicting pharmacoresistance in pediatric epilepsy, *Epilepsia* **54**(Suppl 2) (2013) 19–22.
64. B. Seker Yilmaz, C. Okuyaz and M. Komur, Predictors of intractable childhood epilepsy, *Pediatr. Neurol.* **48**(1) (2013) 52–55.
65. L. G. Kim, T. L. Johnson, A. G. Marson and D. W. Chadwick, Prediction of risk of seizure recurrence after a single seizure and early epilepsy: Further results from the MESS trial, *Lancet Neurol.* **5**(4) (2006) 317–322.
66. F. Mormann, T. Kreuza, C. Rieke, R. G. Andrzejak, A. Kraskov, P. David, C. E. Elgera and K. Lehnertza, On the predictability of epileptic seizures, *Clin. Neurophysiol.* **116**(3) (2005) 569–587.
67. J. E. Walker, Power spectral frequency and coherence abnormalities in patients with intractable epilepsy and their usefulness in long-term remediation of seizures using neurofeedback, *Clin. EEG Neurosci.* **39**(4) (2008) 203–205.
68. M. Nuwer, Assessment of digital EEG, quantitative EEG, and EEG brain mapping: Report of the American Academy of Neurology and the American Clinical Neurophysiology Society, *Neurology* **49**(1) (1997) 277–292.
69. S. K. Kessler, P. R. Gallagher, R. A. Shellhaas, R. R. Clancy and A. G. Bergqvist, Early EEG improvement after ketogenic diet initiation, *Epilepsy Res.* **94**(1–2) (2011) 94–101.
70. C. C. Jouny, G. K. Bergey and P. J. Franaszczuk, Partial seizures are associated with early increases in signal complexity, *Clin. Neurophysiol.* **121**(1) (2010) 7–13.
71. G. M. Tedrus, L. C. Fonseca, E. M. Melo and V. L. Ximenes, Educational problems related to quantitative EEG changes in benign childhood epilepsy with centrotemporal spikes, *Epilepsy Behav.* **15**(4) (2009) 486–490.
72. C. M. Michel, D. Lehmann, B. Henggeler and D. Brandeis, Localization of the sources of EEG delta, theta, alpha and beta frequency bands using the FFT dipole approximation, *Electroencephalogr. Clin. Neurophysiol.* **82**(1) (1992) 38–44.
73. P. L. Nunez, *Quantitative States of Neocortex* (Oxford University Press, London, 1995).
74. P. Gloor, *Electrophysiology of Generalized Epilepsy* (Academic Press, London, 1984).
75. J. C. Sackellares, D. S. Shiau, J. J. Halford, S. M. LaRoche and K. M. Kelly, Quantitative EEG analysis for automated detection of nonconvulsive seizures in intensive care units, *Epilepsy Behav.* **22**(Suppl 1) (2011) S69–S73.