Title

Multimodal data and machine learning for surgery outcome prediction in complicated cases of mesial temporal lobe epilepsy.

Permalink

https://escholarship.org/uc/item/77k4p6ws

Authors

Memarian, Negar

Kim, Sally

Dewar, Sandra

et al.

Publication Date

2015-09-01

DOI

10.1016/j.compbiomed.2015.06.008

Peer reviewed

Multimodal data and machine learning for surgery outcome prediction in complicated cases of mesial temporal lobe epilepsy

Negar Memarian[a,b,*], **Sally Kim**[b], **Sandra Dewar**[c], **Jerome Engel Jr.**[b,c,d,e], and **Richard J. Staba**[b]

[a]Department of Psychology, David Geffen School of Medicine at UCLA, Los Angeles, CA 90095

[b]Department of Neurology, David Geffen School of Medicine at UCLA, Los Angeles, CA 90095

[c]Department of Neurosurgery, David Geffen School of Medicine at UCLA, Los Angeles, CA 90095

[d]Department of Neurobiology, David Geffen School of Medicine at UCLA, Los Angeles, CA 90095

[e]Department of Psychiatry & Biobehavioral Sciences, David Geffen School of Medicine at UCLA, Los Angeles, CA 90095

Abstract

Background—This study sought to predict postsurgical seizure freedom from pre-operative diagnostic test results and clinical information using a rapid automated approach, based on supervised learning methods in patients with drug-resistant focal seizures suspected to begin in temporal lobe.

Method—We applied machine learning, specifically a combination of mutual information-based feature selection and supervised learning classifiers on multimodal data, to predict surgery outcome retrospectively in 20 presurgical patients (13 female; mean age±SD, in years 33±9.7 for females, and 35.3±9.4 for males) who were diagnosed with mesial temporal lobe epilepsy (MTLE) and subsequently underwent standard anteromesial temporal lobectomy. The main advantage of the present work over previous studies is the inclusion of the extent of ipsilateral neocortical gray matter atrophy and spatiotemporal properties of depth electrode-recorded seizures as training features for individual patient surgery planning.

Results—A maximum relevance minimum redundancy (mRMR) feature selector identified the following features as the most informative predictors of postsurgical seizure freedom in this study's sample of patients: family history of epilepsy, ictal EEG onset pattern (positive correlation

_____

[*]Corresponding author: Negar Memarian, Social Cognitive Neuroscience Laboratory, UCLA, Dept. of Psychology, Box 951563, 1285 Franz Hall, Los Angeles, CA 90095-1563, nmemarian@ucla.edu, Phone: 310-739-9509, Fax: 310-825-9257.

with seizure freedom), MRI-based gray matter thickness reduction in the hemisphere ipsilateral to seizure onset, proportion of seizures that first appeared in ipsilateral amygdala to total seizures, age, epilepsy duration, delay in the spread of ipsilateral ictal discharges from site of onset, gender, and number of electrode contacts at seizure onset (negative correlation with seizure freedom). Using these features in combination with a least square support vector machine (LS-SVM) classifier compared to other commonly used classifiers resulted in very high surgical outcome prediction accuracy (95%).

Conclusions—Supervised machine learning using multimodal compared to unimodal data accurately predicted postsurgical outcome in patients with atypical MTLE.

Keywords

Mesial temporal epilepsy; surgical outcome prediction; supervised learning; mutual information

1. Introduction

1.1. Background

Mesial temporal lobe epilepsy (MTLE) is a common type of chronic, medically intractable epilepsy. The recommended treatment for patients with pharmacoresistant MTLE is surgery to remove brain area(s) necessary and sufficient for generating spontaneous seizures, conceptually referred to as the epileptogenic zone (EZ) [1], [2]. However, surgery does not always result in the elimination or complete control of seizures even after patients had standard or tailored resection of the suspected EZ. A review of the literature indicates that a comprehensive diagnostic evaluation that considers results from multiple modalities, for example electrophysiology, neuroimaging, neuropsychological testing, clinical history and semiology, may more accurately delineate the EZ that could improve the likelihood for postsurgical seizure freedom, particularly in cases with complicated MTLE [3], [4], [5], [6], [7], [8], [9]. The inherent differences in brain structure and function of different patients (even those with similar diagnosis), absence of one or more of preoperative data modalities, and lack of a standard multimodal data analysis scheme even when all or majority of these data are available make surgical outcome prediction a nontrivial task. In cases like these, machine learning techniques could be useful because such techniques could perceive obscure associations between multimodal preoperative results and postsurgical outcome in MTLE surgery candidates.

A key advantage of machine learning is the ability to objectively manipulate multimodal, potentially contradictory data, allowing for production of interim results that the algorithm can readily revise as more data becomes available. Machine learning has been used in diverse biomedical applications, such as intelligent patient monitoring in the operating room [10], identifying subtypes of cancer [11], Alzheimer's disease diagnosis [12], and telemedicine [13]. In early diagnosis of Alzheimer's disease, multimodal classifiers that fused EEG, MRI, and PET modalities showed an improvement of up to 20% compared to the classification performance obtained when using each individual data source separately [12].

Machine learning techniques have also been utilized in epilepsy research for automated epilepsy diagnosis [14], [15], [16], seizure lateralization [17], [7], [18], [19], differentiation between mesial and neocortical temporal lobe epilepsies [20], analysis of electrophysiology-hemodynamics connectivity in epileptic patients [21], [22], and in several studies to predict postsurgical seizure freedom [23], [24], [25], [26], [27], [28], [29], [30], [31].

1.2. Purpose

The goal of this study was to computationally evaluate results from preoperative diagnostic tests (intracranial EEG and MRI) and clinical data using supervised learning methods to predict surgical outcome. The premise is to train a computer algorithm by giving it features (preoperative data) along with corresponding labels (surgical outcome score). The computer algorithm tries to find patterns between features that are exclusively associated with the given label. Usually such patterns are not obvious from visual inspection of the data. After the algorithm learns, it will be able to classify new subjects (i.e., predict their surgical outcome) based on features only. The main advantage of the present work over previous studies is the inclusion of the extent of ipsilateral neocortical gray matter atrophy and spatiotemporal properties of depth electrode-recorded seizures as training features for individual patient surgery planning.

1.3. Significance

The proposed method is objective, automated and fast. In the present study, it is used to predict one of the two classes of seizure free or not seizure free outcome in drug resistant epileptic patients. If used in a multi-classes manner (e.g., (1) 0-25%, (2) 25-50% chance, (3) 50-75% chance, (4) 75-100% chance of seizure freedom), the proposed method can quantify the probability of favorable surgery outcome for a (test) patient based on his/her pre-operative multimodal data and the pre-operative and surgical outcome data of previous (training) patients. Also, it can quantify the likelihood of patient membership to various surgical outcome classes. Expandability and capacity to improve accuracy are other advantages of this strategy, as it can learn from previous patients' multimodal data to increase the classification accuracy for new patients iteratively. Ultimately, these methods may provide a basis for a standardized approach for the pre-surgical evaluation of suspected temporal lobe epilepsy or other type of epilepsy as long as the required features are available and can be entered as input to the classifier. Recent work presented at last year's American Epilepsy Society Annual Meeting emphasized a need for a more formal approach in the diagnosis and treatment of non-lesional epilepsy [32]. It would appear that epileptologists need direction on how to approach more difficult cases of epilepsy, including complex cases of MTLE.

The main advantage of the present work compared to previous related studies is the inclusion of the extent and spatial distribution of neocortical atrophy and intracranial electrophysiological recordings as training features [33]. In addition, it is anticipated the proposed strategy could be used to identify patients that might predict a postsurgical outcome other than completely seizure-free, (e.g., improved with 90%, 75%, etc. reduction in seizures), and in these cases, consider the potential benefits from other treatments such as

vagal or trigeminal nerve stimulation or deep brain stimulation. To this end, prognosis data from a pool of patients who underwent non-surgical treatment options is required.

2. Materials and methods

2.1. Subjects

Twenty patients with drug resistant focal seizures suspected to arise from temporal lobe were subjects for this study (13 female). All subjects were candidates for epilepsy surgery, but required presurgical depth electrode to localize the brain area where seizures began because results from non-invasive tests were inconclusive. Depth EEG implantations were bilateral. Mean age ± SD, in years was 33±9.7 for females, and 35.3 ±9.4 for males. All patients gave verbal and written informed consent prior to participating in these research studies, which were approved by the UCLA Office for Protection of Research Subjects Medical Institutional Review Board. The surgery outcome for each patient was labeled as either 1 (seizure free ~ Engel class I with and without auras) or 2 (not seizure free ~ Engel class II or higher) [1]. Table 1 summarizes the demographic and postsurgical outcome information for the study's subjects (epilepsy duration: 22.6±10.5 years, seizure onset age: 11.3±8.8 years, follow up at which outcome classes were determined: 50.3±32.4 months). The electrophysiological and neuroimaging data from patients in the current study were used in a different study that analyzed cortical gray matter thickness in relation to depth electrode-recorded hypersynchronous (HYP) and low voltage fast (LVF) ictal EEG onset patterns [34].

2.2. Feature selection

2.2.1. Features—A combination of 88 demographic, clinical, electrophysiological, and structural magnetic resonance imaging (MRI) features were extracted for each patient as shown in Table 2, which also includes the definition and extraction method for each of these features. These features were selected based on an extensive literature search and included those most commonly cited and used in the presurgical evaluation of drug resistant focal epilepsy (e.g., [1], [3], [8], [35], [36], [37], [38], [39], [40], [41]), as well as consultation with expert epileptologists. The preoperative MRI modality features (features 77 to 80 in Table 2 denote the rate and distribution of gray matter thickness reduction per year of epilepsy duration, when hemispheres ipsilateral to seizure onset were grouped together (features 77 and 79), or when hemispheres contralateral to seizure onset were grouped together (features 78 and 80). GM thickness reduction per each year of epilepsy was determined as the magnitude of the interaction between ictal EEG onset pattern and epilepsy duration on cortical gray matter thickness after controlling for age and gender. For each feature, missing values were replaced by the mean value of that feature column.

2.2.2. Feature selection algorithm options—The purpose of feature selection is to minimize the number of parameters in the final machine learning model to improve performance and generalizability [42]. The central assumption when using a feature selection technique is that the data contains redundant or irrelevant features. Redundant features are those which provide no more information than the currently selected features, and irrelevant features provide no useful information in any context [43]. Feature selection

is also useful as part of the data analysis process, showing which features are important for prediction, and how these features are related [44]. Many feature selection techniques have been created, and generally are grouped into one of three categories: "filter" [45], [46], [47], [44], [48], [49], "wrapper" [50], [51], [52], [53], or "embedded" [54], [55], [56] feature selection methods. The filter approach often selects features by testing whether some preset conditions about the features and the target class are satisfied [49]. For example, these conditions can be measures of distance, information (or uncertainty), dependence, consistency, or classifier error rate. A wrapper is a feature selector that convolves with a classifier (e.g., naïve Bayes classifier) and evaluates the goodness (i.e., classification accuracy) of subset of features being selected by optimizing a criterion (usually minimizing the classification error) [49], [53]. In the embedded feature selection approach, feature selection is built into the inductive algorithm and so the structure of the class of functions under consideration plays a crucial role [55].

2.2.3. Rationale for using the selected feature selection algorithm—Generally, while a wrapper feature selection approach can yield high classification accuracy for a particular classifier, it has the disadvantage of high computational complexity and less generalization of the selected features on other classifiers [49]. The complexity of filter methods is much lower than wrappers, and selected features return comparable classification errors for different classifiers [43], [49]. Between the available filter approaches, the current study used the minimal-redundancy-maximal-relevance algorithm (mRMR) because of its advantages in terms of both feature selection complexity and feature classification accuracy [49]. In this method, relevant features and redundant features are considered simultaneously, i.e., as formulized in formula (1), it seeks to maximize the relevance of a feature set for a specific class and minimize the redundancy of all features in the feature set. Relevance is defined by the average value of all mutual information (MI) values between the individual feature (x_j) and the specific class (c). In formula (1), S_{m-1} is a feature set with $m-1$ features. The task is to select the mth feature from the set $\{X - S_{m-1}\}$. This is done through an incremental search method by selecting the feature that maximizes the condition inside the square brackets. Redundancy is the average value of all MI values (denoted with I in formula (1)) between the individual feature and every other feature it the set (x_i). Furthermore, the mRMR algorithm is suitable for unprocessed data, where the features selected in this way will have more or less correlation with each other. This is because mRMR does not intend to select features that are independent of each other. Instead, at each step, it tries to select a feature that minimizes the redundancy and maximizes the relevance [49].

$$\max_{x_j \in X - S_{m-1}} \left[I\left(x_j;c\right) - \frac{1}{m-1} \sum_{x_i \in S_{m-1}} I\left(x_j;x_i\right) \right] \quad (1)$$

The interpretation of the mRMR output (ranked features) may be better clarified with an example. Consider a small three variable data set, for which the mRMR output has ranked the features from highest to lowest as follows: 2, 1, and 3. This result means that the combination of features 2 and 1 is better than the combination of 2 and 3. It also means 2 is the best feature if you only want one feature, and both "2 and 1" are the best combination if

you want two features. However, this does not mean 3 is the least relevant or most redundant feature. If features are selected without considering the relationship between all the features, but only between individual features and the target class variable, results may even show that 3 is more relevant than 1. However the combination of 2 and 1 is better than that of 2 and 3 [57]. mRMR has been tested on a number of medical datasets, namely HDR MultiFeat [58], Arrhythmia [58], NCI [59], [60], and Lymphoma [61], and led to promising improvement on classification accuracy. These studies involved datasets with large number for features, combination of discrete and continuous data, and different classifiers. Considering that our dataset has similar attributes, we were inspired to use the mRMR feature selection method for the present study.

2.3. Classifiers

Five widely used supervised learning classifiers [62], [63], [64] were tested in this study: (1) Linear Discriminant Analysis (LDA), (2) Naïve Bayes (NB), (3) support vector machines (SVM) with radial basis function kernel (SVM-rbf), (4) SVM with multilayer perceptron kernel (SVM-mlp), and (5) Least-Square SVM (LS-SVM). One or more of these classifiers were utilized in previous studies on supervised learning based prediction of MTLE surgery outcome [23], [24], [25], [26], and we chose these classifiers to build up on previous work and show the potential superiority of our proposed scheme. The normality of data was checked for the LDA and NB algorithms. The common premise of all these classifiers is to train a computer algorithm by giving it features (preoperative data) along with corresponding labels (surgery outcome score). The computer algorithm tries to find patterns between features that are exclusively associated with the given label. Usually such patterns are not obvious from manual inspection of the data. After the algorithm learns, it will be able to classify a new subject (i.e., predict his/her surgery outcome) based on his/her preoperative features only. A leave-one-out cross-validation scheme was used to train and test each classifier. In other words for each classifier, the prediction accuracy was computed as the average of twenty iterations (as there were twenty subjects), where at each iteration one patient was left out as the test subject and the remaining nineteen patients were used for training the classifier. This procedure was repeated twenty times until every patient was used as a test subject. It should be noted that feature selection was performed only on the training data, i.e., feature selection repeated in each leave one out iteration.

In the case of the LS-SVM classifier, we tested the prediction accuracy of the classifier with three different kernels, namely linear, polynomial, and RBF. To verify the optimal range of parameter values for the kernel function, we repeated the prediction task 1000 times and computed the histograms for the average regularization parameter and average kernel bandwidth over those iterations.

2.3.1. Convergence criterion—In order to select the optimal number of features and classification accuracy for each classifier, an iterative approach was implemented. For all classifiers, classification started with one feature (top ranked feature determined from mRMR), and in each iteration, another feature (in the order of ranking) was added. This process terminated, when the following criteria were satisfied:

$$(C_{i-2} \geq C_{i-3}) \ \& \ (C_{i-1} \geq C_{i-2}) \ \& \ (C_{i-1} \geq C_i) \quad \text{(2)}$$

where C is the classification accuracy, and i is the number of features used. Upon satisfaction of this criterion, $(i-1)$ and C_{i-1} were selected as the optimal number of features and optimal classification accuracy respectively.

3. Results

In a related, but separate study, we measured cortical gray matter (GM) thickness in a group of patients diagnosed with unilateral MTLE that had either depth electrode-recorded HYP (hypersynchronous) or LVF (low voltage fast) ictal EEG onset patterns [34]. These MRI-based measures of cortical GM thickness and electrophysiological data, as well as clinical information obtained from medical histories, were used as data in the current analysis. Table 3 provides a list of the ten most informative features determined by the mRMR feature selection algorithm. According to Table 3, family history of epilepsy, the maximum ipsilateral GM thickness reduction per each year of epilepsy, and the ratio of the number of seizures that first appeared in the ipsilateral amygdala to total number of seizures were collectively the strongest predictors of surgical outcome in this dataset. Table 3 also lists the extent and direction of impact of each of these features on postsurgical outcome. Outcome was one of 1 = seizure free, 2= not seizure free. A positive correlation indicates that an increase in the state of the feature correlates with an increase in the state of the postsurgical outcome, while a negative value indicates an increase in the state of the feature correlates with a decline in the state of the postsurgical outcome. For example, patients with a family history of epilepsy (feature state 1 = family history of epilepsy 'yes', see Table 2, last column, feature #84) were more likely to have seizure free postsurgical outcome (outcome state 1 = seizure free with or without aura).

Figure 1 shows the prediction accuracy of each of the five classifiers versus number of features selected for training. Each graph converges to the maximum prediction accuracy for that classifier. The optimal number of features for each classifier was found based on equation (1) and correspond to features 1-5 in Table 3. Three of the classifiers (LDA, NB, and SVM-mlp) converged to their optimal prediction accuracy (80%, 80%, and 85% respectively) with the smallest number of features for training (i.e., 3 features). LS-SVM required the greatest number of training features (i.e., 5 features) to reach its optimal prediction accuracy of 95%. We tested the significance of this result by running a Mann-Whitney U test between prediction accuracy values (20 values for the 20 iterations of the cross-validation scheme) computed by the LS-SVM classifier trained on our data and prediction by chance. The results indicated a statistically significant difference between the predication accuracy of our prediction method from chance at 5% significance level (p=0.0082). However, using a similar approach it was found that at their optimal prediction accuracy, the performance of the tested classifiers was not statistically significant from each other (p-values > 0.05 for all possible pairs of classifiers).

Prediction accuracy for each of the classifiers in relation to the different feature categories is shown in Figure 2. In this analysis, all features belonging in one category were used to train

the classifiers. There were a total of four runs, one for each of the four feature categories, i.e., electrophysiology, MR imaging, clinical history, and demographic data. LDA and NB prediction accuracies could not be calculated with electrophysiological features only because the pooled covariance of the training data was not positive definite. The highest prediction accuracy for an individual feature category was achieved using the SVM-rbf classifier with MRI data (80%), whereas only 25% prediction accuracy was found using the SVM-mlp classifier with these same MRI data, which was the lowest value of any classifier and feature category. With respect to feature category, clinical features were associated with the highest average prediction accuracy across all classifiers (mean: 65%), while the lowest was found with MRI data and an average accuracy of 56%. The LS-SVM classifier did the best or nearly the best with each of the feature categories (mean: 70%).

The prediction accuracy of the LS-SVM classifier with various kernels is presented in Table 4. The RBF kernel resulted in the highest prediction accuracy (95%). Figure 3a shows a sample histogram of the optimal regularization parameter estimated by the LS-SVM model with this kernel (RBF) across the twenty cross validation iterations (because there were twenty subjects). It can be seen that the most frequent estimate of the regularization parameter was in the range of 0 to 5. Histogram of the optimal RBF kernel bandwidth estimated by the LS-SVM model across the twenty cross validation iterations is depicted in Figure 3b, with the interval 5 to 15 being the most frequent estimated range of this parameter. A larger kernel bandwidth indicates a stronger smoothing. Figures 3a and 3b show the histograms for the average regularization parameter and average kernel bandwidth over 1000 repeated iterations of the prediction task. The optimal range of parameter values for the kernel function is further confirmed from this repeated test. The average value of the regularization parameter estimated by the LS-SVM model fell in the 0 to 5 interval in 91% of the iterations (914 out of 1000). The average value of the estimated kernel bandwidth fell in the 5 to 15 interval in 99.4% of the iterations (994 out of 1000).

4. Discussion

Results from the current study showed that a machine learning-based approach using data from multiple modalities could predict with high accuracy postsurgical seizure freedom in patients with atypical unilateral MTLE. Furthermore, the proposed approach using a combination of intracranial EEG, quantitative MRI, and clinical data did a significantly better job of predicting outcome than machine learning using intracranial EEG (Phase II) or imaging (MRI SPM) alone.

Between the investigated classifiers, results suggest LS-SVM is the most promising based on its high classification accuracy using a small number of features. Machine learning models that achieve similar accuracy by operating on a selected set of features are preferred in investigative research over machine learning models that are saturated with input features [42]. LS-SVM enjoys the sparseness of SVM solutions but is also more computationally efficient, i.e., is not limited by the time and memory consumed optimization [65]. Our results confirm the results of a previous study that showed the power of LS-SVM for the purpose of surgical outcome prediction in TLE [23]. Using a data fusion approach and a set of eleven features for each subject that study succeeded to prospectively predict that surgery

was the best solution for a patient in more than 90% of the cases. The [23] study examined pre-operative interictal and ictal scalp EEG findings, age of onset, gender, duration of epilepsy, risk factors, family history, and physical resected tissue as features, but did not discuss the strength of each feature in predicting postsurgical outcome. A recent study also used supervised classification data mining to predict surgical outcome from common clinical features and neuropsychological and pathological evaluations in a group of 23 MTLE patients with HS [25]. The analysis revealed that outcome could be predicted with an estimated accuracy of almost 90% using some clinical and neuropsychological features, which confirms the promise of machine learning for surgery outcome prediction. It should be noted that the aforementioned study was specific to MTLE patients with unilateral HS only. In our study patients included those with and without HS and a broader range of features (most importantly quantitative structural MRI measurements) were used that could be informative in difficult cases of MTLE.

An advantage of the SVM-based methods is their flexibility to customize the classifier's kernel function and regularization parameter to achieve optimum results for the data in hand. The task of fine-tuning the SVM-based models is instrumental for one to achieve optimal performance [66]. Because of its superior performance compared to the rest of the classifiers, we focused on fine-tuning the LS-SVM method to optimize prediction accuracy. The LS-SVM algorithm computes the optimal regularization parameter such that the complexity of the model is minimized, while good fitting of the training data points is achieved. It can be seen from Figure 3a that the most frequent estimate of the regularization parameter was in the range of 0 to 5. This agrees with the fact that SVM models with low values of regularization parameter tend in general to achieve better performance than those with high values of this parameter [66]. As depicted in Figure 3b, the interval 5 to 15 was the most frequent estimated range for the optimal RBF kernel bandwidth. A larger kernel bandwidth indicates a stronger smoothing. For investigative purposes, we repeated the prediction task 1000 times and computed the histograms for the average regularization parameter and average kernel bandwidth over those iterations (Figures 3(c) and 3(d)). The optimal range of parameter values for the kernel function was further confirmed from this repeated test. The average value of the regularization parameter estimated by the LS-SVM model fell in the 0 to 5 interval in 91% of the iterations (914 out of 1000). The average value of the estimated kernel bandwidth fell in the 5 to 15 interval in 99.4% of the iterations (994 out of 1000).

To the best of our knowledge, this is the first study that included cortical GM thickness measurements computed through quantitative structural MRI analysis for individual MTLE patient surgical outcome prediction [33]. The information associated with these MRI data likely contributed to the higher prediction accuracy because evidence from MRI studies suggests that the presence of extra-temporal lobe lesions can reduced the likelihood for postsurgical seizure freedom [67], [8], [68], [69], [70]. Quantitative MRI techniques have been used to investigate changes in hippocampal [71], [72], [73] or extra-hippocampal [74], [75], [76] GM thickness between patient groups and controls, or between different epilepsy patient groups. However, since results from these studies were based on group averaged differences, it is difficult to extend these results to surgical planning in individual patients. By utilizing machine learning, we have presented an approach to use quantitative MRI

derived from individual subjects. To this end, the extent and spatial distribution of cortical atrophy compared to age and gender matched controls, computed using a sophisticated quantitative cortical pattern matching technique [77] were employed among the input features to the supervised learning classifiers. Maximum cortical GM thickness reduction per each year of epilepsy ipsilateral to hemisphere of onset proved to be a very informative predictor of surgical outcome (ranked 2 by mRMR; Table 3). It can be seen from Figure 2 that the MRI features alone yield an accuracy of 80% with SVM-rbf classifier, which further signifies the importance of these of features in surgical outcome prediction.

Presurgical intracranial studies are needed when non-invasive studies, i.e., scalp EEG, MRI, PET, SPECT, are inconclusive. Another novelty of this study was inclusion of features from patients' intracranial EEG recordings for post surgery outcome prediction. Five of the top 10 selected features by the mRMR algorithm (proportion of seizures to total seizures that first appeared in the ipsilateral amygdala and anterior region of middle temporal gyrus, ictal EEG onset pattern, delay in seconds from initial seizure activity to appearance at other recording sites in ipsilateral hemisphere, and number of contacts at seizure onset), belong to the set of intracranial EEG (electrophysiological) features. Based on these findings, amygdala was more predictive than any other mesial structure. While electrophysiological features alone predicted the correct surgical outcome in 50% to 70% of the cases (Figure 2; SVM-rbf: 70%, SVM-mlp: 50%, LS-SVM: 70%), together with their preceding ranked features, they yielded a classification accuracy of 85% or higher with the SVM classifiers (Figure 1; SVM-rbf: 90%, SVM-mlp: 85%, LS-SVM: 95%). These results highlight the added value in employing electrophysiological features in addition to the imaging and clinical features and corroborate the intracranial EEG procedure carried out in some major hospitals as part of the preoperative evaluation protocol. In patients who have an MRI that indicates extra-temporal lobe abnormalities, electrophysiological studies are carried out to confirm the epileptogenicity of a lesion [71]. In some cases, intracranial EEG recordings are required to directly record from structural abnormalities that could exist on mesial and lateral surfaces of each cerebral hemisphere. Deep brain electrical activity recorded through implantation of intracranial electrodes reveals important information about the sites of seizures, number of seizures at each site, onset pattern of seizures, and the latency of seizure propagation from onset site to other ipsilateral and contralateral structures. While phase II recordings can be very informative, it is not comprehensive as it suffers from the limitation of discrete spatial coverage. i.e., depth electrodes are not implanted in all brain areas, and even this partial implant configuration is not consistent among all patients. Moreover, a potential drawback of the proposed approach is that it may not be comparably effective without the availability of the intracranial EEG recordings, which is in fact not collected in all hospitals. The association of greater number of contacts at seizure onset with a higher chance of seizure freedom post surgery, as well as the association of greater number of seizures recorded in the ipsilateral amygdale with a higher chance of seizure freedom post surgery confirms the benefit of amygdalohippocampectomy to this patient population.

Other studies have pointed to the importance of seizure onset pattern in surgical outcome prediction. In the present study, the onset pattern was one of the two most common MTLE seizure onset patterns, namely hypersynchronous (HYP) and low voltage fast (LVF). As reflected in Table 1, seizure-free outcome was more likely in patients with HYP ictal onset

(89% in HYP patients, 64% in LVF). In another study, we found that a larger proportion of HYP than low voltage fast (LVF) onset seizures terminate with EEG suppression or irregular slowing [34]. Another potential reason for better post surgical outcome in patients with HYP ictal onset pattern might be the presence of non-epileptogenic areas of atrophy. While this group had more diffuse GM atrophy compared to the LVF group, the atrophic areas are not necessarily the source of epilepsy. It has been shown that extra temporal lesions on MRI are not always the source of epileptic activity and may simply be incidental and unrelated to the seizure [78].

Our results suggest that some clinical factors play a critical role in the success of epilepsy surgery. Family history of epilepsy was selected as the top rank feature by mRMR. Presence of family history of MTLE correlates with a higher chance of seizure freedom post surgery (Table 3). This is in agreement with previous studies that reported family history of epilepsy as a predictor of remission [79], [80], [81], [82]. Another important clinical feature was epilepsy duration. In a study to quantify the epileptogenicity of brain structures recorded with depth electrodes, Bartolomei et al found a statistically significant correlation between the duration of epilepsy and the number of structures disclosing high epileptogenicity, suggesting that MTLE is a gradually evolving process in which the epileptogenicity of the temporal lobe tends to increase with time [83]. Our results indicate a statistically insignificant negative correlation between epilepsy duration and post surgical seizure freedom (Table 3), implying that longer duration of epilepsy prior to surgery was correlated with a higher chance of seizure freedom post surgery. While this counters the results reported by [84] that favorable long-term surgical outcome was associated with shorter epilepsy duration, it is more in line with more recent studies that found no significant associations between post surgical outcome and duration of epilepsy [85], [86], [87]. The discrepancy could be due to the variable length of the follow-up across studies, as the chance of seizure remission after surgery can be significantly affected by when the follow-up assessment is performed [3]. Moreover, comparing the columns "age" (patient's age at time of surgery) and epilepsy duration "number of years of epilepsy prior to surgery" in Table 1 shows that nearly all of the patients had epilepsy for half of their life and 12 of them (out of 20), had epilepsy for over two thirds of their life. This reflects the clinical institution's willingness to operate on older patients, and could be because this was an atypical population of patients, with longstanding epilepsy who did not consider surgery earlier because it was relatively mild, while the other patients may have had earlier surgery because they had more severe epilepsy and were desperate.

In the present study, mean age at time of surgery of patients who achieved seizure freedom (Engel class I) at least 47 months after surgery was older than the mean age of those patients whose follow-up indicated an Engel class II or higher. Previous studies pointed to early age of seizure onset as a risk factor for poor prognosis after surgery [41], [88], [9]. In the present study, age at seizure onset was not available for all patients, and hence the impact of this feature on postsurgical outcome prediction could not be verified.

While the correlation coefficients listed in Table 3 are useful in revealing the direction of the association of each feature and postsurgical outcome, the small value of the coefficients as well as the rather large p-values give the impression that the strength of each feature-

outcome correlation was not impressive. In spite of this, through the use of supervised learning classifiers, the combination of these features predicted postsurgical outcome very successfully. This further highlights the importance of multimodal machine learning in extracting associations and patterns that are not obvious from visual inspection of the data.

It should be noted that subjects of the present study were all patients who were suspected to have temporal lobe epilepsy and thus surgical candidates. Depth electrode studies localized the seizure onset zone to mesial temporal structures, which in most cases had MRI signal abnormalities, and all had anteromesial temporal lobectomy [89]. The question whether this approach works in individuals with neocortical epilepsy remains to be investigated.

5. Future Work

This is a retrospective study, with a selected of number MTLE surgery candidates. To make definitive statements and potentially use it in clinical practice, a prospective study (have the algorithm predict outcome before surgery) with a larger population seems necessary. As for any supervised learning scheme, a larger training sample space (ideally 88 subjects, i.e., the number of features, or more) is expected to improve the classification accuracy. We expect this improvement to be more significant for the linear classifiers (LDA and Naïve Bayes) compared to the SVM classifiers, as the latter group already yielded a very high prediction accuracy of 90% or higher, even with the present smaller sample space. The algorithms can be extended to address multi-class classification (e.g., prediction of post surgical Engel class instead of 'seizure free' or 'not seizure free' outcome).

Weights can be given to features based on the significance of their effect on epilepsy experts' decision making about surgery. This can be tailored to the specific profile of each case. For example, in a patient with extra-temporal spread damage, even more emphasis should be put on depth electrode EEG recordings to increase the likelihood of a favorable surgery outcome. The assignment of weights can be objectively based on the information content (mRMR ranking) of the features or manually through expert clinician's input.

Upon availability of required facilities in health care centers, new features can be added to the feature space, which in turn may enhance the outcome prediction accuracy further. Such features include but not limited to scalp EEG (this might be a difficult feature since there is not a consensus on how it is interpreted among different centers), Positron Emission Tomography (PET) results, magnetoencephalography (MEG) results, seizure semiology and patient's behavioral features during seizures, and seizure etiology (e.g., head trauma, infections such as encephalitis). The latter two features may be less reliable, since in the absence of detailed medical reports, they are either subjective or based on the recollection of patient or his/her family. FDG-PET and MRI are common in presurgical evaluation and provide important information in diagnosis of MTLE. Typically, when patients are PET and MRI negative, i.e., subtle or no evidence for interictal temporal lobe hypometabolism or HS respectively, they are further evaluated through depth electrode studies. These features should certainly be added to the feature space in future analyses.

Inclusion of more comprehensive neuropsychological test results or a continuous rather than discrete spectrum of neuropsychological features may further enhance the prediction

accuracy of the classifiers. It has been suggested that personality style features from Rorschach test results are linked with post surgical outcome in epilepsy patients, such that the majority (85%) of extroversive (EB2) patients had successful (Engel I) outcome [25].

The patients in the present study were atypical cases of MTLE, which are difficult and often require invasive studies for surgical planning. In our future studies, we plan to select an equal number of patients who were, and were not, seizure free, and also include patients who did not go to surgery. Reliably predicting who will fail phase II, might save a lot of money, discomfort, and risk.

6. Conclusion

Utilization of multimodal data and machine learning can result in high accuracy surgical outcome prediction in patients with mesial temporal lobe epilepsy, which is not attainable from unimodal assessments. A combination of maximum relevance minimum redundancy (mRMR) feature selector and least square support vector machine (LS-SVM) classifier, resulted in best surgery outcome prediction accuracy. Based on the mRMR ranking, family history of epilepsy, rate of gray matter thickness reduction in the hemisphere ipsilateral to seizure onset, proportion of seizures that first appeared in ipsilateral amygdala to total seizures, age, epilepsy duration, seizure onset pattern, latency from initial seizure activity to appearance at sites in ipsilateral hemisphere, gender, and number of electrode contacts at seizure onset were the most informative predictors of surgical outcome.

Acknowledgments

References

1. Engel, J, Jr.. Seizures and Epilepsy Contemporary Neurology Series. Oxford University Press; NY, USA: 2013.

2. Engel J Jr. et al. Early surgical therapy for drug-resistant temporal lobe epilepsy: a randomized trial. JAMA. 2012; 307(9):922–30. [PubMed: 22396514]

3. Tonini C, et al. Predictors of epilepsy surgery outcome: a meta-analysis. Epilepsy Res. 2004; 62(1): 75–87. [PubMed: 15519134]

4. Hauptman JS, Salamon N, Mathern GW. Neuroimaging in the definition and organization of the epilepsies: we're not there yet. Epilepsia. 2012; 53(Suppl 2):22–7. [PubMed: 22765501]

5. French JA, et al. Characteristics of medial temporal lobe epilepsy: I. Results of history and physical examination. Ann Neurol. 1993; 34(6):774–80. [PubMed: 8250525]

6. Williamson PD, et al. Characteristics of medial temporal lobe epilepsy: II. Interictal and ictal scalp electroencephalography, neuropsychological testing, neuroimaging, surgical results, and pathology. Ann Neurol. 1993; 34(6):781–7. [PubMed: 8250526]

7. Moser DJ, et al. Electroencephalographic, volumetric, and neuropsychological indicators of seizure focus lateralization in temporal lobe epilepsy. Arch Neurol. 2000; 57(5):707–12. [PubMed: 10815137]

8. Thom M, et al. Mesial temporal lobe epilepsy: How do we improve surgical outcome? Ann Neurol. 2010; 68(4):424–34. [PubMed: 20976764]

9. Blumcke I, et al. A new clinico-pathological classification system for mesial temporal sclerosis. Acta Neuropathol. 2007; 113(3):235–44. [PubMed: 17221203]

10. Jones, RW. Data Fusion in the O.R.: Improved Clinical Measurements and Intelligent Patient Monitoring. 1st International Conference on Sensing Technology; Palmerston North, New Zealand. 2005; p. 326-331.

11. Yuan Y, Savage RS, Markowetz F. Patient-specific data fusion defines prognostic cancer subtypes. PLoS Comput Biol. 2011; 7(10):e1002227. [PubMed: 22028636]

12. Polikar R, et al. Multimodal EEG, MRI and PET data fusion for Alzheimer's disease diagnosis. Conf Proc IEEE Eng Med Biol Soc. 2010; 2010:6058–61. [PubMed: 21097123]

13. Megalooikonomou V, Kontos D. Medical data fusion for telemedicine. A model for distributed analysis of medical image data across clinical information repositories. IEEE Eng Med Biol Mag. 2007; 26(5):36–42. [PubMed: 17941321]

14. Kerr WT, et al. Automated diagnosis of epilepsy using EEG power spectrum. Epilepsia. 2012; 53(11):e189–92. [PubMed: 22967005]

15. Shoeb, AH.; Guttag, JV. ICML. Omnipress; Haifa, Isreal: 2010. Application of Machine Learning To Epileptic Seizure Detection; p. 975-982.

16. Isik H, Sezer E. Diagnosis of epilepsy from electroencephalography signals using multilayer perceptron and Elman Artificial Neural Networks and Wavelet Transform. J Med Syst. 2012; 36(1):1–13. [PubMed: 20703754]

17. Kerr WT, et al. Computer-Aided Diagnosis and Localization of Lateralized Temporal Lobe Epilepsy Using Interictal FDG-PET. Front Neurol. 2013; 4:31. [PubMed: 23565107]

18. Keihaninejad S, et al. Classification and lateralization of temporal lobe epilepsies with and without hippocampal atrophy based on whole-brain automatic MRI segmentation. PLoS One. 2012; 7(4):e33096. [PubMed: 22523539]

19. Fakhraei, S.; Soltanianzadeh, H.; Elisevich, K.; Fotouhi, F. Attribute Ranking for Lateralizing Focal Epileptogenicity in Temporal Lobe Epilepsy. Proceedings of the 17th Iranian Conference of Biomedical Engineering (ICBME2010); Iran. 2010; p. 1-4.

20. Pfander M, et al. Clinical features and EEG findings differentiating mesial from neocortical temporal lobe epilepsy. Epileptic Disord. 2002; 4(3):189–95. [PubMed: 12446221]

21. Gotman J, et al. Combining EEG and fMRI: a multimodal tool for epilepsy research. J Magn Reson Imaging. 2006; 23(6):906–20. [PubMed: 16649203]

22. Laufs H, et al. Recent advances in recording electrophysiological data simultaneously with magnetic resonance imaging. Neuroimage. 2008; 40(2):515–28. [PubMed: 18201910]

23. Rezaie, MG.; Soltanian-Zadeh, H.; Siadat, MR.; Elisevich, KV. Soft computing approaches to computer aided decision making for temporal lobe epilepsy.. Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS); Ann Arbor, Michigan, USA. 2005; IEEE;

24. Fakhraei, S.; Soltanianzadeh, H.; Jafari-Khouzani, K.; Elisevich, K.; Fotouhi, F. Confident surgical decision making in temporal lobe epilepsy by heterogeneous classifier ensembles. In: Myra Spiliopoulou, HW.; Cook, Diane; Pei, Jian; Wang, Wei; Zaïane, Osmar; Wu, Xindong, editors. IEEE 11th International Conference on Data Mining Workshops (ICDMW). IEEE; Vancouver, Canada: 2011. p. 1003-1009.

25. Armananzas R, et al. Machine learning approach for the outcome prediction of temporal lobe epilepsy surgery. PLoS One. 2013; 8(4):e62819. [PubMed: 23646148]

26. Feis DL, et al. Prediction of post-surgical seizure outcome in left mesial temporal lobe epilepsy. Neuroimage Clin. 2013; 2:903–11. [PubMed: 24179841]

27. Bernhardt BC, H.S. Bernasconi A, Bernasconi N. Magnetic Resonance Imaging Pattern Learning in Temporal Lobe Epilepsy: Classification and Prognostics. Annals of Neurology. 2015; 77(3): 436–446. [PubMed: 25546153]

28. Yankam Njiwa Y, G.K. Costes N, Mauguiere F, Ryvlin P, Hammers A. Advanced [18F]FDG and [11C] flumazenil PET analysis for individual outcome prediction after temporal lobe epilepsy surgery for hippocampal sclerosis. NeuroImage: Clinical. 2015; 7:122–131. [PubMed: 25610774]

29. Armañanzas R, A.-N.L. Defelipe-Oroquieta J, Kastanauskaite A, de Sola RG, Defelipe J, Bielza C, Larrañaga P. Machine learning approach for the outcome prediction of temporal lobe epilepsy surgery. PLoS One. 2013; 8(4):e6281.

30. Arle JE, P.K. Devinsky O, Doyle WK. Neural network analysis of preoperative variables and outcome in epilepsy surgery. J Neurosurg. 1999; 90(6):998–1004. [PubMed: 10350243]

31. Grigsby J, K.R. Schneiders JL, Gates JR. Brewster Smith W, Predicting outcome of anterior temporal lobectomy using simulated neural networks. Epilepsia. 1998; 39(1):61–66. [PubMed: 9578014]

32. Trinka, E.; Wilfong, A. Future Perspectives in the Management of Refractory SE. American Epilepsy Society (AES) 66th Annual Meeting; San Diego, CA.. Dec 2012;

33. Memarian N, et al. Quantitative analysis of structural neuroimaging of mesial temporal lobe epilepsy. Imaging Med. 2013; 5(3) 10.2217/iim.13.28.

34. Memarian N, et al. Ictal Depth EEG and MRI Structural Evidence for Two Different Epileptogenic Networks in Mesial Temporal Lobe Epilepsy. PLoS ONE. 2015; 104:e0123588. [PubMed: 25849340]

35. Baldauf CM, et al. Surgical outcome in patients with refractory epilepsy associated to MRI-defined unilateral mesial temporal sclerosis. Arq Neuropsiquiatr. 2006; 64(2B):363–8. [PubMed: 16917602]

36. Pittau F, et al. Prognostic factors in patients with mesial temporal lobe epilepsy. Epilepsia. 2009; 50(Suppl 1):41–4. [PubMed: 19125847]

37. Schulz R, et al. Analysis of reoperation in mesial temporal lobe epilepsy with hippocampal sclerosis. Neurosurgery. 2011; 68(1):89–97. discussion 97. [PubMed: 21099715]

38. Smyth MD, et al. Outcome following surgery for temporal lobe epilepsy with hippocampal involvement in preadolescent children: emphasis on mesial temporal sclerosis. J Neurosurg. 2007; 106(3 Suppl):205–10. [PubMed: 17465386]

39. Keller SS, et al. Persistent seizures following left temporal lobe surgery are associated with posterior and bilateral structural and functional brain abnormalities. Epilepsy Res. 2007; 74(2-3): 131–9. [PubMed: 17412561]

40. Varoglu AO, et al. Prognosis of patients with mesial temporal lobe epilepsy due to hippocampal sclerosis. Epilepsy Res. 2009; 85(2-3):206–11. [PubMed: 19345070]

41. Aguglia U, et al. Age at onset predicts good seizure outcome in sporadic non-lesional and mesial temporal sclerosis based temporal lobe epilepsy. J Neurol Neurosurg Psychiatry. 2011; 82(5):555–9. [PubMed: 20971758]

42. Kerr WT, et al. The utility of data-driven feature selection: Re: Chu et al. 2012. Neuroimage. 2013

43. Langley, P. Selection of relevant features in machine learning. AAAI Fall Symposium on Relevance; LA. 1994; New Orleans, USA: AAAI Press; p. 127-131.

44. Kira, K.; Rendell, L. A practical approach to feature selection.. Ninth International Conference on Machine Learning; Morgan Kaufmann: Aberdeen, Scotland. 1992; p. 249-256.

45. Almuallim, H.; Dietterich, TG. Learning with many irrelevant features. Ninth National Conference on Artificial Intelligence; 1991; San Jose, CA.: AAAI Press; p. 547-552.

46. Cardie, C. Using decision trees to improve case-based learning. Tenth International Conference on Machine Learning; Morgan Kaufmann: Amherst, MA. 1993; p. 28-36.

47. Doak, J. An evaluation of feature-selection methods and their application to computer security. University of California Davis; Davis, CA.: 1992.

48. Kubat, M.; Flotzinger, D.; Pfurtscheller, G. Discovering patterns in EEG signals: Comparative study of a few methods. European Conference on Machine Learning; 1993; Vienna, Austria: Springer-Verlag; p. 367-371.

49. Peng H, Long F, Ding C. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. IEEE Trans Pattern Anal Mach Intell. 2005; 27(8):1226–38. [PubMed: 16119262]

50. Langley, P.; Sage, S. AAAI-94 Workshop on Case-Based Reasoning. AAAI Press; Seattle, WA.: 1994. Oblivious decision trees and abstract cases; p. 113-117.

51. Moore, AW.; Lee, MS. Eleventh International Conference on Machine Learning. Morgan Kaufmann; New Brunswick, N J.: 1994. Efficient algorithms for minimizing cross validation error; p. 190-198.

52. Townsend-Weber, T.; Kibler, D. AAAI94 Workshop on Case-Based Reasoning. AAAI Press; Seattle, WA.: 1994. Instancebased prediction of continuous value; p. 30-35.

53. John, GH.; Kohavi, R.; Pfleger, K. Eleventh International Conference on Machine Learning. Morgan Kaufmann; New Brunswick, NJ.: 1994. Irrelevant features and the subset selection problem; p. 121-129.

54. Maldonado S, W.R. Embedded Feature Selection for Support Vector Machines: State-of-the-Art and Future Challenges. Lecture Notes on Computer Science. 2011; 7042:304–311.

55. Saeys Y II, Larranaga P. A review of feature selection techniques in bioinformatics. Bioinformatics. 2007; 23(19):2507–2517. [PubMed: 17720704]

56. Monteiro, ST.; Murphy, RJ. Embedded feature selection of hyper spectral bands with boosted decision trees. IEEE International Geoscience and Remote Sensing Symposium (IGARSS); 2011; Vancouver, BC, Canada: IEEE; p. 2361-2364.

57. Ding, C.; Peng, H. Minimum redundancy feature selection from microarray gene expression data. 2nd IEEE Computational Systems Bioinformatics Conference (CSB 2003); Aug 2003; Stanford, CA.: IEEE; p. 523-528.

58. Ausuncion, A.; Newman, D. UCI Machine Learning Repository. UC Irvine; 2007.

59. Ross DT, et al. Systematic variation in gene expression patterns in human cancer cell lines. Nat Genet. 2000; 24(3):227–35. [PubMed: 10700174]

60. Scherf U, et al. A gene expression database for the molecular pharmacology of cancer. Nat Genet. 2000; 24(3):236–44. [PubMed: 10700175]

61. Alizadeh AA, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature. 2000; 403(6769):503–11. [PubMed: 10676951]

62. Duda, RO.; Hart, PE.; Stork, DG. Pattern classification. 2 ed.. John Wiley & Sons; USA: 2012.

63. Drugowitsch, J. Studies in Computational Intelligence. Springer-Verlag; Berlin, Germany: 2008. Design and analysis of learning classifier systems A prbabilistic approach..

64. Soman, KP.; Loganathan, R.; Ajay, V. Machine Learning with SVM and Other Kernel Methods. PHI Learning Pvt. Ltd; New Delhi, India: 2009.

65. Wang, H.; Hu, D. Comparison of SVM and LS-SVM for Regression. International Conference on Neural Networks and Brain (ICNN&B '05); Oct, 2005; Beijing, China: IEEE; p. 279-283.

66. Lima CA, Coelho AL, Eisencraft M. Tackling EEG signal classification with least squares support vector machines: a sensitivity analysis study. Comput Biol Med. 2010; 40(8):705–14. [PubMed: 20621291]

67. Bonilha L, et al. Subtypes of medial temporal lobe epilepsy: Influence on temporal lobectomy outcomes? Epilepsia. 2012; 53(1):1–6. [PubMed: 22050314]

68. Eross L, et al. Interhemispheric propagation of seizures in mesial temporal lobe epilepsy. Ideggyogy Sz. 2009; 62(9-10):319–25. [PubMed: 19835273]

69. Mueller SG, et al. Voxel-based optimized morphometry (VBM) of gray and white matter in temporal lobe epilepsy (TLE) with and without mesial temporal sclerosis. Epilepsia. 2006; 47(5): 900–7. [PubMed: 16686655]

70. Cohen-Gadol AA, et al. Long-term outcome of epilepsy surgery among 399 patients with nonlesional seizure foci including mesial temporal lobe sclerosis. J Neurosurg. 2006; 104(4):513–24. [PubMed: 16619654]

71. Ogren JA, et al. Three-dimensional hippocampal atrophy maps distinguish two common temporal lobe seizure-onset patterns. Epilepsia. 2009; 50(6):1361–70. [PubMed: 19054395]

72. McDonald CR, et al. Subcortical and cerebellar atrophy in mesial temporal lobe epilepsy revealed by automatic segmentation. Epilepsy Res. 2008; 79(2-3):130–8. [PubMed: 18359198]

73. Keller SS, et al. Voxel based morphometry of grey matter abnormalities in patients with medically intractable temporal lobe epilepsy: effects of side of seizure onset and epilepsy duration. J Neurol Neurosurg Psychiatry. 2002; 73(6):648–55. [PubMed: 12438464]

74. Lin JJ, et al. Reduced neocortical thickness and complexity mapped in mesial temporal lobe epilepsy with hippocampal sclerosis. Cereb Cortex. 2007; 17(9):2007–18. [PubMed: 17088374]

75. Keller SS, Roberts N. Voxel-based morphometry of temporal lobe epilepsy: an introduction and review of the literature. Epilepsia. 2008; 49(5):741–57. [PubMed: 18177358]

76. Bernhardt BC, et al. Mapping thalamocortical network pathology in temporal lobe epilepsy. Neurology. 2012; 78(2):129–36. [PubMed: 22205759]

77. Thompson PM, et al. Mapping cortical change in Alzheimer's disease, brain development, and schizophrenia. Neuroimage. 2004; 23(Suppl 1):S2–18. [PubMed: 15501091]

78. Alsaadi TM, et al. Potentially misleading extratemporal lobe lesions in patients with temporal lobe epilepsy. J Neurol Neurosurg Psychiatry. 2003; 74(5):566–9. [PubMed: 12700292]

79. Andermann, E. Genetic aspects of epilepsy. In: Robb, P., editor. Epilepsy updates: causes and treatment. Symposia Specialists; Miami, FL.: 1980. p. 11-24.

80. Andermann, E. Multifactorial inheritance of generalized and focal epilepsy. In: A. V.E., et al., editors. Genetic basis of the epilepsies. Raven Press; New York, NY: 1982. p. 355-374.

81. Kobayashi E, et al. Outcome of surgical treatment in familial mesial temporal lobe epilepsy. Epilepsia. 2003; 44(8):1080–4. [PubMed: 12887440]

82. Elsharkawy AE, et al. Long-term outcome after temporal lobe epilepsy surgery in 434 consecutive adult patients. J Neurosurg. 2009; 110(6):1135–46. [PubMed: 19025359]

83. Bartolomei F, Chauvel P, Wendling F. Epileptogenicity of brain structures in human temporal lobe epilepsy: a quantified study from intracerebral EEG. Brain. 2008; 131(Pt 7):1818–30. [PubMed: 18556663]

84. Janszky J, et al. Temporal lobe epilepsy with hippocampal sclerosis: predictors for long-term surgical outcome. Brain. 2005; 128(Pt 2):395–404. [PubMed: 15634733]

85. Lowe NM, et al. The duration of temporal lobe epilepsy and seizure outcome after epilepsy surgery. Seizure. 2010; 19(5):261–3. [PubMed: 20430656]

86. Krsek P, et al. Predictors of seizure-free outcome after epilepsy surgery for pediatric tuberous sclerosis complex. Epilepsia. 2013; 54(11):1913–21. [PubMed: 24117179]

87. Jehi LE, et al. Temporal lobe epilepsy surgery failures: predictors of seizure recurrence, yield of reevaluation, and outcome following reoperation. J Neurosurg. 2010; 113(6):1186–94. [PubMed: 20831360]

88. Clusmann H, et al. Lesional mesial temporal lobe epilepsy and limited resections: prognostic factors and outcome. J Neurol Neurosurg Psychiatry. 2004; 75(11):1589–96. [PubMed: 15489392]

89. Fried I. Magnetic resonance imaging and epilepsy: neurosurgical decision making. Magn Reson Imaging. 1995; 13(8):1163–70. [PubMed: 8750331]

Summary

Mesial temporal lobe epilepsy (MTLE) is characterized by epileptic seizures that begin in or primarily involve mesial temporal limbic structures and is frequently associated with focal structural abnormalities, such as hippocampal sclerosis, which can be detected on neuroimaging. The recommended treatment for patients with drug-resistant MTLE is surgery to remove brain area(s) necessary and sufficient for generating spontaneous seizures, conceptually referred to as the epileptogenic zone (EZ). However, surgery does not always result in the elimination or complete control of seizures even after patients had standard or tailored resection of the suspected EZ. A review of the literature indicates that a comprehensive diagnostic evaluation that considers results from multiple modalities, for example electrophysiology, neuroimaging, neuropsychological testing, clinical history and semiology, may more accurately delineate the EZ that could improve the likelihood for postsurgical seizure freedom, particularly in cases with complicated MTLE. The inherent differences in brain structure and function of different patients (even those with similar diagnosis), absence of one or more of preoperative data modalities, and lack of a standard multimodal data analysis scheme even when all or majority of these data are available make surgical outcome prediction a non-trivial task. In cases like these, machine learning techniques could be useful because such techniques could perceive obscure associations between multimodal preoperative results and postsurgical outcome in MTLE surgery candidates. In this paper we computationally evaluate results from preoperative diagnostic tests and clinical data using supervised learning methods to predict surgical outcome. These methods may provide a basis for a standardized approach for the pre-surgical evaluation of suspected temporal lobe epilepsy or other type of epilepsy as long as the required features are available and can be entered as input to the classifier. A combination of maximum relevance minimum redundancy (mRMR) feature selector and least square support vector machine (LS-SVM) classifier, resulted in very high surgery outcome prediction accuracy (95%). Based on the mRMR ranking, family history of epilepsy, rate of gray matter thickness reduction in the hemisphere ipsilateral to seizure onset, proportion of seizures that first appeared in ipsilateral amygdala to total seizures, age, epilepsy duration, seizure onset pattern, latency from initial seizure activity to appearance at sites in ipsilateral hemisphere, gender, and number of electrode contacts at seizure onset were the most informative predictors of surgical outcome.

Highlights

- Machine learning with multimodal data can accurately predict postsurgical outcome in patients with drug resistant mesial temporal lobe epilepsy.

- Features resulting from quantitative analysis of structural MRI and intracranial EEG are informative predictors of postsurgical outcome.

- Least-square support vector machine with radial basis function kernel resulted in optimal prediction.

- Clinical factors such as family history of epilepsy and duration of epilepsy significantly affect the chance of seizure freedom post epilepsy surgery.
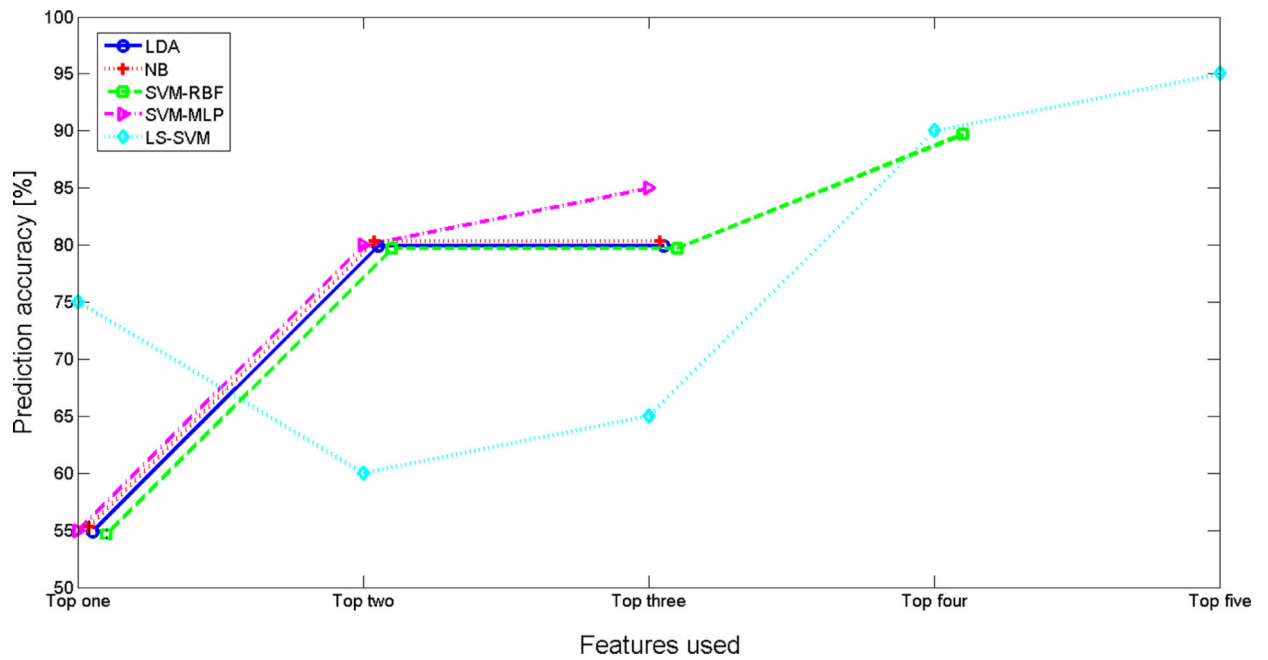
Figure 1.
Prediction accuracy of each of LDA, NB, SVM-rbf, SVM-mlp and LS-SVM classifiers versus number of features selected by mRMR for training (refer to Table 3). The optimal number of features for each classifier was found based on equation (1). Each graph converges to the maximum prediction accuracy for that classifier.
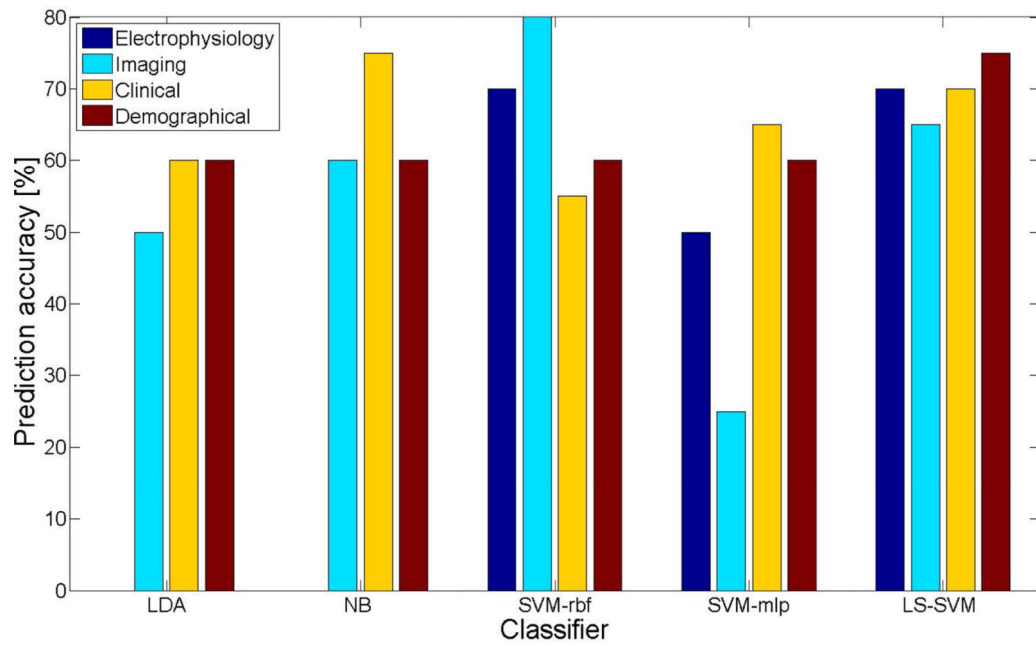
Figure 2.
Prediction accuracy of the classifiers when trained with one feature category only. For every classifier, each bar represents the prediction accuracy for one features category, namely electrophysiological (dark blue), MRI (light blue), clinical (yellow), and demographical (red). LDA and NB prediction accuracies could not be calculated with electrophysiological features only, because the pooled covariance of the training data was not positive definite.
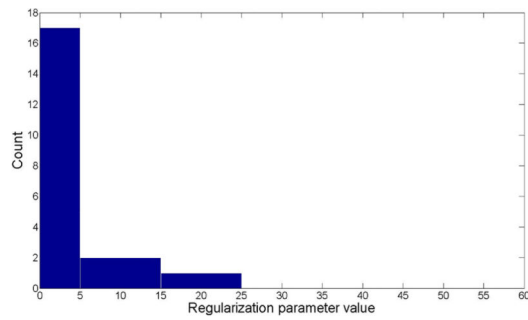
(a)
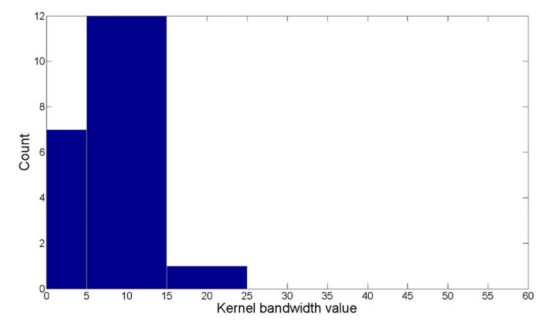
(b)
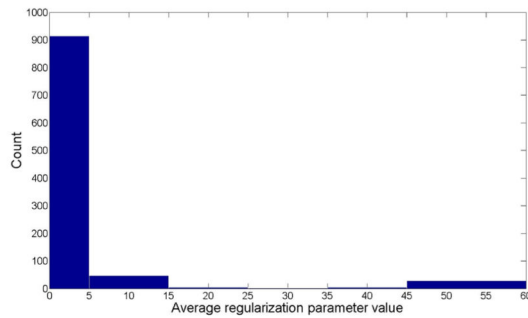
(c)

(d)

Figure 3.
A sample histogram of the optimal (a) regularization parameter, and (b) RBF kernel bandwidth estimated by the LS-SVM model (with the RBF kernel) across the twenty cross validation iterations. Panels (c) and (d) depict histogram of the average regularization parameter, and RBF kernel bandwidth estimated by the LS-SVM model (with the RBF kernel) over 1000 iterations, respectively.

Summary of study subjects.

Table 1

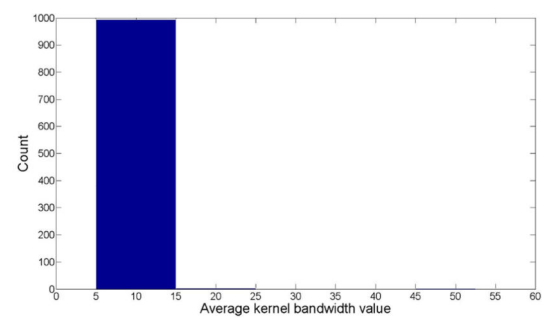| Patient	Hemi	Pattern	Contacts onset	Contacts ipsi	Contacts cont	Lat ipsi	Lat contra	GM thin ipsi	GM lobe ipsi	GM thin cont	GM lobe cont	HS	Dur	Freq	Fam	Psych	Febrile	Gender	Age	Follow-up	Outcome ¥
1	R	HYP	1	1	1	13.5	58.5	1.29	2	1.58	1	No	24	3.5	No	Yes	No	M	31.2	36	SF
2	L	HYP	3	3	6	14	58	1.43	2	1.34	1	Yes	30	2	No	Yes	Yes	F	42	51	SF
3	R	HYP	1	4	2.5	10	59	1.5	1	1.54	2	No	20	27	Yes	Yes	No	F	39	36	SF
4	R	HYP	1	1	5	2	128	0.85	2	0.68	2	No	20	150.5	Yes	No	No	F	25	60	SF
5	L	HYP	2	5.5	2	5	54	18	2	0.99	2	No	20.4	1	Yes	Yes	Yes	F	22	72	SF
6	R	HYP	3	4	9	18	31	1	2	0.95	1	No	30	8	No	Yes	No	F	38	32	SF
7	L	HYP	2	2.5	3	5	16	26	1	1.07	2	No	20	27	Yes	No	Yes	F	40	45	SF
8	R	HYP	3	7	12	14	32	21	1	1.26	2	No	12	12	No	Yes	No	M	25	60	SF
9	R	LVF	3	0	3	0	6	42	1	1.58	1	Yes	34	4	Yes	Yes	No	F	37.8	79	SF
10	R	LVF	9	1.5	3	0.3	2	14	1	1.32	2	Yes	32	4.5	Yes	Yes	Yes	M	38	38	SF
11	L	LVF	2	6.5	3	4	48.5	33	2	1.32	2	Yes	41	16	No	No	No	M	44	120	SF
12	L	LVF	3	1	2	15	28	1.5	2	1.22	1	No	20	1.5	No	No	No	M	41	24	SF
13	L	LVF	9	5	4	35	40.8	1	1	0.93	2	Yes	33	64	No	Yes	Yes	F	47	9	SF
14	L	LVF	2	9	4	32	49	1	2	1.41	2	Yes	6	8	No	Yes	No	M	46	12	SF
15	L	LVF	5	0	3	10.4	24	39	2	1.35	2	Yes	23	2	No	Yes	No	F	29.8	23	SF
16	R	HYP	3	4.5	4	8	40.8	1.2	2	1.14	1	No	9	30	No	Yes	No	F	18	126	NSF
17	L	LVF	1	1	2	9	40	71	1	1.37	2	No	40	2.5	No	Yes	Yes	F	43	47	NSF
18	R	LVF	2	2	3.5	5	154	0.65	2	0.78	1	No	20	2.5	No	Yes	No	F	28	27	NSF
19	R	LVF	5	3	2	4	13.5	1	2	1.03	2	Yes	9	6	No	No	No	M	22	84	NSF
20	L	LVF	9	6	3	3.5	28	0.75	2	1.41	2	Yes	8	5	No	No	No	F	20	24	NSF

Hemi: Hemisphere of seizure onset, Pattern: Ictal EEG onset pattern, Contacts onset: No. of contacts at seizure onset, Contacts ipsi: No. of contacts ipsilateral spread, Contact cont: No. of contacts contralateral spread, Lat ipsi: Latency ipsilateral spread (sec), Lat contra: Latency contralateral spread (sec), GM thin ipsi: GM thickness reduction ipsilateral (mm per year), GM lobe ipsi: No. of lobes affected with GM thickness reduction ipsilateral, GM thin cont: GM thickness reduction contralateral (mm per year), GM lobe cont: No. of lobes affected with GM thickness reduction contralateral, HS: Hippocampal sclerosis, Dur: Epilepsy duration (years), Freq: Seizure frequency (per month), Fam: Family history, Psych: Neuropsychological test results, Febrile: History of febrile seizures, Gender: female (F) or male (M), Age: age at time of surgery, Follow-up: number of months after surgery, when the follow-up assessment was performed. Outcome: Post surgery outcome.

¥
At an average follow-up of 54 months.

Table 2

List of features extracted for each patient.

Category	No.	Feature	Definition	Modality	Representation
Electrophysiology	1	Hemisphere of seizure onset	Hemisphere where seizures begin	Intracranial EEG	Binary (1 = L, 2 = R)
	2	Ictal EEG onset pattern	HYP: high amplitude, low frequency (<=2 Hz) spike or spike-and-wave discharges lasting 5 sec or longer LVF: low amplitude, high frequency (10 Hz) activity superimposed or not on EEG slow wave	Intracranial EEG	Binary (1 = HYP, 2 = LVF)
	3	No. of contacts at seizure onset	Number of electrode contacts where seizure activity first appeared	Intracranial EEG	Continuous
	4	No. of contacts ipsilateral spread	Number of contacts from initial seizure activity to appearance at sites in ipsilateral hemisphere	Intracranial EEG	Continuous
	5	No. of contacts contralateral spread	Number of contacts from initial seizure activity to appearance at sites in contralateral hemisphere	Intracranial EEG	Continuous
	6	Latency ipsilateral spread (sec)	Delay in seconds from initial seizure activity to appearance at sites in ipsilateral hemisphere	Intracranial EEG	Continuous
	7	Latency contralateral spread (sec)	Delay in seconds from initial seizure activity to appearance at sites in contralateral hemisphere	Intracranial EEG	Continuous
	8 to 56	Anatomical recording site(s) where seizure first appeared, including ipsilateral amygdala; entorhinal cortex; anterior, middle, and posterior hippocampus; posterior parahippocampal gyrus; anterior, middle, posterior regions of inferior, middle, and superior temporal gyrus; medial, middle, and lateral orbitofrontal cortex, anterior cingulate gyrus, and supplementary motor cortex	Proportion of seizures that began at each ipsilateral recording site divided by the total number of seizures	Intracranial EEG	Continuous
	47 to 76	Anatomical recording site(s) where seizure first appeared, including contralateral amygdala; entorhinal cortex; anterior, middle, and posterior hippocampus; posterior parahippocampal gyrus; anterior, middle, posterior regions of inferior, middle, and superior temporal gyrus; medial, middle, and lateral orbitofrontal cortex, anterior cingulate gyrus, and supplementary motor cortex	Proportion of seizures that began at each contralateral recording site divided by the total number of seizures	Intracranial EEG	Continuous

Category	No.	Feature	Definition	Modality	Representation
Imaging	77	GM thickness reduction ipsilateral (mm per year)	Maximum GM thickness reduction per each year of epilepsy ipsilateral	Preoperative MRI	Continuous
	78	No. of lobes affected with GM thickness reduction ipsilateral	Number of lobes affected with significant GM thickness reduction compared to controls ($p<0.05$) ipsilateral	Preoperative MRI	Binary (1 = one lobe affected, 2 = more than one lobe affected)
	79	GM thickness reduction contralateral (mm per year)	Maximum GM thickness reduction per each year of epilepsy contralateral	Preoperative MRI	Continuous
	80	No. of lobes affected with GM thickness reduction contralateral	Number of lobes affected with significant GM thickness reduction compared to controls ($p<0.05$) contralateral	Preoperative MRI	Binary (1 = one lobe affected, 2 = more than one lobe affected)
	81	Hippocampal sclerosis	Presence of HS on MRI	Preoperative MRI	Binary (1 = yes, 2 = no)
Clinical	82	Epilepsy duration (years)	Number of years of epilepsy prior to surgery	Patients' medical charts	Continuous
	83	Seizure frequency (per month)	Number of seizures per month	Patients' medical charts	Continuous
	84	Family history	Familial history of epilepsy	Patients' medical charts	Binary (1 = yes, 2 = no)
	85	Neuropsychological test results	Abnormal neuropsychological test results or psychiatric history (e.g., depression, insomnia, family background)	Patients' medical charts	Binary (1 = yes, 2 = no)
	86	Febrile seizure	History of febrile seizures	Patients' medical charts	Binary (1 = yes, 2 = no)
Demographical	87	Gender	-	Patients' medical charts	Binary (1 = female, 2 = male)
	88	Age	Patient's age at time of surgery	Patients' medical charts	Continuous

Table 3

Top ten features linked to postsurgical outcome based on the mRMR feature selection algorithm and their correlation with postsurgical denoted as 1 (seizure free) or 2 (not seizure free) For each feature, r is the correlation coefficient between two vectors, the first vector being the value of that feature across all subjects and the second vector being the postsurgical outcome (seizure free or not seizure free) across all subjects.

Rank	Feature number[¥]	Feature name	r	p	Interpretation of r-value with respect to seizure-free outcome (1)
1	84	Family history	0.38	0.10	Presence of family history of epilepsy
2	77	GM thickness reduction ipsilateral (in mm per year)	−0.35	0.13	Ipsilateral reduction in cortical GM thickness per year of epilepsy
3	8	Amygdala mesial ipsi	−0.31	0.18	Greater ratio of number of seizures recorded in ipsilateral amygdala to total number of seizures recorded
4	88	Age	−0.35	0.13	Older patient age at time of surgery
5	82	Epilepsy duration	−0.35	0.13	Longer duration of epilepsy prior to surgery
6	2	Seizure onset pattern	0.29	0.21	HYP ictal EEG onset pattern
7	6	Latency ipsilateral spread (sec)	−0.35	0.13	Longer delay in ipsilateral spread of ictal discharges from site(s) of onset
8	10	Amygdala lateral ipsi	−0.13	0.59	Greater number of seizures recorded in the lateral aspect of anterior temporal lobe
9	87	Gender	−0.18	0.44	Male patient
10	3	No. of contacts at seizure onset	−0.20	0.40	Greater number of electrode contacts involved at seizure onset

p is the p-value for testing the hypothesis of no correlation against the alternative that there is a nonzero correlation. The continuous features were discretized based on a median split.

[¥]This number is with reference to feature numbers in Table 2.

Table 4

Postsurgical outcome prediction accuracy of the LS-SVM classifier with various kernel functions.

LS-SVM kernel function	Prediction accuracy
Linear	85%
Polynomial	90%
RBF	95%