Using Machine Learning to Detect Mesial Temporal Lobe Epilepsy Based on Structural and Functional Neuroimaging

Baiwan Zhou[1], Dongmei An[2], Fenglai Xiao[2,3], Running Niu[1], Wenbin Li[1], Wei Li[2], Xin Tong[2], Graham J Kemp[4], Dong Zhou[2] *, Qiyong Gong[1,5], Du Lei[1,6,7] *

[1] Huaxi MR Research Center (HMRRC), Department of Radiology, West China Hospital of Sichuan University, Chengdu, Sichuan, China

[2] Department of Neurology, West China Hospital of Sichuan University, Chengdu, Sichuan, China

[3] Department of Clinical and Experimental Epilepsy, Institute of Neurology, University College London, London, UK

[4] Institute of Ageing and Chronic Disease, Faculty of Health and Life Sciences, University of Liverpool, Liverpool, UK

[5] Department of Psychology, School of Public Administration, Sichuan University, Chengdu, Sichuan, China

[6] Department of Psychosis Studies, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

[7] Department of Psychiatry and Behavioral Neuroscience, University of Cincinnati, Cincinnati, Ohio, USA

* Corresponding Author: Du Lei, leidu@ucmail.uc.edu or Dong Zhou, zhoudong66@yahoo.de

ABSTRACT

Mesial temporal lobe epilepsy (mTLE), the commonest type of focal epilepsy, is associated with both functional and structural brain alterations. Recently, machine learning (ML) techniques have been successfully used to discriminate mTLE from healthy controls. However, most have used either functional or structural neuroimaging data as input, without exploiting the opportunity to combine both. We conducted a multimodal ML study based on both functional and structural neuroimaging measures. We enrolled 37 left mTLE and 37 right mTLE patients and 74 healthy controls, and trained a support vector machine learning model to distinguish them using each single measure as well as their combinations. For each single measure, we obtained a mean accuracy of 74% and 70% for discriminating left mTLE and right mTLE from controls, respectively, and 65% putting all patients together. For left mTLE, we achieved an accuracy of 78% using functional data and 79% using structural data, while the highest accuracy of 84% was obtained when combining all functional and structural measures. These findings suggest that combining multi-modal measures within a single model could be a promising direction for improving the classification of individual patients with mTLE.

Keywords: mesial temporal lobe epilepsy; functional magnetic resonance imaging; structural magnetic resonance imaging; machine learning; support vector machine.

Introduction

Mesial temporal lobe epilepsy (mTLE) is the commonest type of focal epilepsy in adults, and its pathophysiological substrate is usually hippocampal sclerosis (HS) [1]. Magnetic resonance imaging (MRI) methods have come to play a pivotal role in the evaluation of patients with mTLE, notably resting state functional MRI (rs-fMRI) [2, 3] and structural MRI (sMRI) [4, 5]. However, previous MRI studies have typically measured average group-level differences, rather than evaluating individual patients.

In biological neurology there is growing interest in the application of machine learning (ML) techniques to neuroimaging data for the diagnosis of epilepsy [6-8], and mTLE has been the main focus of this work [9, 10]. Most previous studies using ML techniques to investigate mTLE have used a single neuroimaging modality: multiparameter sMRI data was found to discriminate mTLE patients from controls with 81% accuracy [11], while combining 6 rs-fMRI measures achieved 83% accuracy [12]. In addition, ML has been applied in this way to other neuroimaging modalities such as diffusion tensor imaging (DTI) [13, 14].

Although these studies have shown the potential of integrating appropriate MRI data with ML to detect mTLE with acceptable accuracy, questions remain. Firstly, brain alteration in left mTLE is reportedly more extensive than in right mTLE [15-17], which suggests different patterns of brain abnormalities. However, it is not known whether detecting right and left mTLE separately is better than pooling them when using ML. Secondly, although both structural and functional brain abnormalities have been reported in mTLE [18, 19], no study has so far integrated both structural and functional

data with ML, an approach which improves classification performance in some other diseases [20].

We therefore set out to use ML to distinguish mTLE patients and healthy controls (HC) by combining both sMRI and rs-fMRI data. As inputs for classification, from sMRI we extracted 3 measures which have been successfully used in the investigation of mTLE: gray matter (GM) [21] and white matter (WM) [22] density and cortical thickness [23]; from rs-fMRI the inputs were amplitude of low frequency fluctuation (ALFF) [24] and regional homogeneity (ReHo) [25, 26]. These 5 measures were combined to provide integrated information on the functional and structural brain alterations in mTLE.

We hypothesized (i) that detecting right mTLE and left mTLE patients separately, rather than pooling, would improve classification accuracy; (ii) that the combination of structural and functional measures within a multi-modal, multi-measure model would yield more accurate classification; and (iii) that temporal lobe would contribute the most to the classification, being the main site of abnormality in mTLE.

Materials and methods

Participants

From September 2013 to January 2018, 74 mTLE patients were consecutively recruited in the Department of Neurology in West China Hospital of Sichuan University (Chengdu, China), all of whom met the International League Against Epilepsy criteria for diagnosis of mTLE [27, 28] and were right-handed. All patients had

unilateral HS (37 left and 37 right), as assessed by hippocampal atrophy on T1-weighted MRI (qualitative assessments by two radiologist) and increased signal on T2 fluid-attenuated inverted recovery (FLAIR) in the mesial temporal region. Video EEG was used to confirm that seizure onset was in the ipsilateral temporal lobe. No other mass brain lesion, traumatic brain injury or any psychiatric disorder was apparent in the MRI, EEG and neuropsychological examination. In addition, 74 age-matched, sex-matched right-handed HC were enrolled, all free of any neurological or psychiatric disorders at the time of the study. Table 1 gives the demographic and clinical characteristics of the study groups.

This study was approved by the West China Hospital Clinical Trials and Biomedical Ethics Committee of Sichuan University, and written informed consent was obtained from all of the participants. The study protocol was performed in accordance with the approved guidelines.

Table 1. Demographic and clinical data [a]

Variables	Left mTLE	Right mTLE	HC
Sample size	37	37	74
Age [b]	24.4±8.0	25.2±7.3	25.9±7.8
Gender (M/F)	18 (48.6%)/ 19 (51.4%)	19 (51.4%)/ 18 (48.6%)	37 (50%)/ 37 (50%)
Disease duration (years) [b]	10.4±9.0	12.4±6.9	-
Onset of epilepsy (years) [c]	14.6±9.1	14.3±8.3	-
Initial precipitating insults:			
febrile seizures	11 (30.0%)	8 (21.6%)	
CNS infection	6 (16.2%)	7 (18.9%)	
status epilepticus	None	None	
Antiepileptic drugs:			
monotherapy	10 (27.0%)	9 (24.3%)	-
ditherapy	14 (37.9%)	12 (32.4%)	
multiple AEDs	13 (35.1%)	16 (43.3%)	

[a] Data are presented as mean±standard deviation. No significant differences were identified between the groups in age and gender.
[b] Age and duration of episode were defined at the time of MRI scanning.
[c] This is the age at onset of seizures or diagnosis of mTLE.
Abbreviations: mTLE, mesial temporal lobe epilepsy; HC, healthy controls; AEDs, antiepileptic drugs.

MRI Data Acquisition

MRI scanning was performed with a 3 T system (Tim Trio; Siemens Healthineers, Erlangen, Germany) using an 8-channel phased array head coil. Each functional examination contained 200 image volumes, and total imaging time 410 seconds. Participants were instructed not to focus their thoughts on anything in particular and to keep their eyes closed during the acquisition. Head motion was minimized by using foam pads. Functional scanning parameters: repetition/echo time 2000/30 ms; flip angle 90°; 30 axial sections per volume; 5 mm section thickness (no gap); 64×64

matrix; field of view 240×240 mm^2; voxel size 3.75×3.75×5 mm^3. Structural scanning used a spoiled gradient-recalled sequence to obtain high-resolution three-dimensional T1-weighted images. Structural scanning parameters: 176 slices; slice thickness 1 mm; flip angle 9°; matrix size 256 × 256; repetition/inversion/echo time 1900/900/2.26 ms; voxel size 1 ×1×1 mm^3.

MRI data analysis

Figure 1 shows an overview of the classification approach. Five individual measures were analyzed using a Support Vector Machine (SVM) learning model: cortical thickness, GM and WM, extracted from sMRI data, and ALFF and ReHo, extracted from rs-fMRI data.

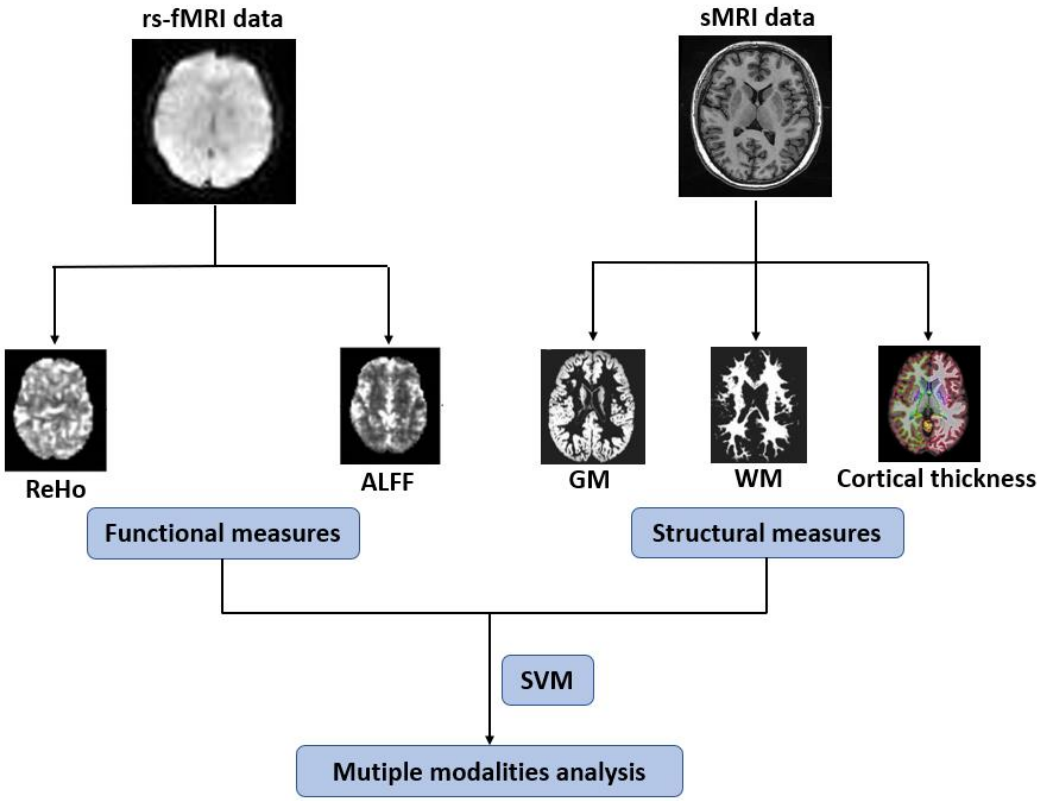

Figure 1. Overview of the classification approach used to assess the diagnostic value of sMRI and rs-fMRI data. Abbreviations: sMRI, structural MRI; rs-fMRI, resting state functional MRI; GM, gray matter; WM, white matter; ReHo, regional homogeneity; ALFF, amplitude of low-frequency fluctuation

Functional data pre-processing

Rs-fMRI data was preprocessed using the DPARSF 4.3 Advanced Edition (http://www.restfmri.net) as follows: removal of the first 10 volumes of each subject's rest data to minimize the impact of instability in the initial MRI signal; correction for acquisition delay between slices; regression of white matter nuisance signals, cerebral spinal fluid blood oxygen level dependent signal using Friston 24 head-motion profiles and scrubbing regressors to minimize the effect of head motion;

normalization using EPI templates (voxel size 3×3×3 mm); smoothing with a Gaussian kernel of 4 mm full-width at half-maximum (FWHM); finally, filtering functional data (band pass: 0.01–0.1 Hz) to reduce the effects of low-frequency drift, and high-frequency noise smoothing.

ReHo maps were then extracted from the pre-processed images using DPARSF software. After removing linear trends in the unsmoothed images and applying a band-pass filter ($0.01 < f < 0.08$ Hz) to reduce low-frequency drift and high-frequency respiratory and cardiac noise, ReHo maps were generated by calculating the concordance of Kendall's coefficient (values from 0 to 1) of the time series of a given voxel with those of its 26 nearest neighbors. The ReHo value of each voxel was standardized by dividing it by the global (within-brain) mean ReHo value.

The ALFF was calculated using DPARSF software. After application of a band-pass filter (0.01–0.08 Hz) and removal of linear trends, the time series was transformed to the frequency domain using fast Fourier transforms. The square root of the power spectrum was calculated and averaged across 0.01–0.08 Hz for each voxel to yield the ALFF. The ALFF of each voxel was standardized by dividing it by the global (within brain) mean ALFF value.

Structural data pre-processing

The 3D T1-weighted images were preprocessed using the Diffeomorphic Anatomical Registration Through Exponentiated Lie (DARTEL) toolbox based on SPM8

(http://www.fil.ion.ucl.ac.uk/spm) as follows: the structural image was segmented into GM and WM; anatomical registration was performed using DARTEL algebra in SPM8 for registration, normalization and modulation; the registered images were transformed to Montreal Neurological Institute (MNI) space (voxel size 1.5×1.5×1.5 mm); the normalized, non-modulated images (GM and WM density images) were smoothed with a 10 mm full-width at half-maximum Gaussian kernel to increase the signal to noise ratio. The preprocessed GM and WM probability maps (the density of GM and WM were reflected by voxel density) were used as measures for the ML analysis.

Cortical thickness was calculated using FreeSurfer software (http://surfer.nmr.mgh.harvard.edu/). The 3D T1-weighted images were processed with the recon-all processing pipeline for cortical reconstruction and volumetric segmentation [29]; the streamlined pipeline included the removal of non-brain tissue, Tailarach transformations, segmentation of subcortical white and deep gray matter regions, intensity normalization and atlas registration. A mesh model of the cortical surface was generated, and the cortical surface was parcellated into 34 cortical regions based on gyral and sulcal landmarks for each hemisphere according to the Desikan–Killiany atlas [30]; cortical thickness for each of these 34 cortical regions were calculated per hemisphere. To improve the ability to detect population changes, we blurred each participant's morphometric parameter map using a 25 mm full-width at half-maximum surface-based Gaussian kernel. Finally, we combined the cortical thickness maps of the left and right hemisphere into a whole brain map,

and this was used as a measure for the ML analysis.

Machine learning classification and evaluation of models

We used SVM [31] to perform single-subject classification. SVM maps the input vectors to a feature space using a set of mathematical functions known as kernels. In this space the model finds the optimum separation surface that maximizes the margin between different classes within a training dataset. Once the separation surface is determined, it can be used to predict the class of new observations using an independent testing dataset. Here a linear kernel was preferred to a nonlinear one to minimize the risk of overfitting. The model was based on LIBSVM [32] and implemented by the Scikit-Learn library [33].

To investigate the performance of each SVM model, we used a 10-fold stratified cross-validation approach. The participants were first divided into 10 non-overlapping partitions, each partition maintaining the same ratio of mTLE patients to HC as the whole group. In each iteration, one partition was considered as the independent test set (where the performance metric is calculated), and the remaining subjects were defined as the training sample. Within each training set, we performed an internal cross-validation (i.e. 10-fold stratified nested cross-validation) to select the optimal set of hyperparameters of the ML models. The linear SVM has only one hyperparameter (the soft margin parameter C) that controls the trade-off between reducing training errors and having a larger separation margin. This parameter was optimized by performing a grid search in the following values: C = 10^{-3}, 10^{-2}, 10^{-1}, 10^{0}, 10^{1}, 10^{2}, 10^{3}, 10^{4}. This yielded the optimum C value for each

input measure. The set of parameters which performed best across the internal cross-validations was selected for each imaging modality and used to train the SVM models.

During multiple measure analysis we combined the SVM predictions of single measures using a weight averaging method (soft voting), which is reportedly slightly more effective than either sum of kernels or multi-kernel learning [34]. We first trained each SVM using a single measure; this allowed us to estimate the likelihood of an individual belonging to the patient or control group (using the Scikit-Learn library default method). Next we calculated the weight probabilities of each specific measure by multiplying its predicted probabilities by a coefficient optimized. After the grid searches for the C parameter, a second nested cross-validation was performed to optimize the coefficient of each specific measure for the soft voting. Each coefficient was evaluated using a grid search with a coefficient search space assuming an integer value between 1 to 10. This second nested cross-validation was also performed using a 10-fold stratified cross-validation. In both nested cross-validations, the highest mean balanced accuracy (defined as the mean of sensitivity and specificity) of the model was used to define the best hyperparameter value. Sensitivity and specificity are taken into account simultaneously by using mean balanced accuracy to optimize the model, which is better than simple accuracy when the samples of two group are unbalanced. Finally, we calculated the average of the predicted weight probabilities, which are the weighted averages of the probability that the SVM model based on each measure predicted that an individual

subject belong to the two groups respectively, the group with the highest score being defined as the predicted class for a given subject.

After the training of SVM models, the final step was to evaluate the performance of the SVM model in conjunction with the evaluation data. To avoid the influence of imbalanced datasets in left or right mTLE analysis, we calculated the balanced accuracy for each SVM models. We also report the sensitivity, specificity, recall, F_1 score and area under the receiver operating characteristic curve (AUC) to evaluate the performance more comprehensively. To obtain meaningful confidence intervals and p-values for each cluster, a random permutation test (1000 times) was used to examine the statistical significance of the classification models.

Creation of discriminative brain region maps

An anatomical automatic labeling (AAL) atlas consisting of 90 regions of interest (ROIs) was used to construct maps of discriminative brain regions [35]. The weight maps are the spatial representation of the decision function that defined the level of each ROI's contributions to the classification process. We report the top 10 discriminative regions of four measures (ReHo, ALFF, GM, WM), in order to seek objective biomarkers of mTLE. The regions that were in the top 10 discriminative regions for more than two measures were defined as the most discriminative regions. Then we extracted the submaps of these regions based on each measure (ReHo, ALFF, GM, WM) and using an SVM technique based on integration of these submaps to verify the results.

Results

Classification performance

The balanced accuracies, sensitivities, specificities, and p-values for the single-subject classification of patients and HC are reported in Table 2, and Figure 2 shows an overview of the classification accuracy. In the identification of all patients versus HC, we obtained an accuracy of 63% for ReHo, 63% for ALFF, 58% for GM, 72% for WM and 63% for cortical thickness. Dividing the patients into left and right mTLE, for the discrimination of left mTLE from HC we obtained an accuracy of 75% for ReHo, 75% for ALFF, 73% for GM, 76% for WM and 72% for cortical thickness; for the discrimination of right mTLE from HC we obtained an accuracy of 68% for ReHo, 73% for ALFF, 66% for GM, 73% for WM and 66% for cortical thickness. Thus dividing the patients into left and right mTLE allows more accurate classification than pooling all patients, for both left mTLE (mean 74% versus 64%; paired t-test, $p = 0.005$) and right mTLE (mean 69% versus 64%; paired t-test, $p = 0.030$).

Discriminating left mTLE from HC, combining functional measures (ReHo and ALFF) yielded an accuracy of 78%. Combining structural measures (GM, WM and cortical thickness) yielded an accuracy of 79% (Table 2). Thus for both structural and functional modalities, combining different measures yields a marginally higher accuracy of classification than using single measures alone. For discriminating either

all patients or right mTLE from HC we found no such increase in accuracy.

Combining all measures across structural and functional modalities yields an accuracy of 84% for discriminating left mTLE from HC (Table 2), which is higher than the use of single modalities, either structural (84% vs 79%) or functional (84% vs 78%). For discriminating either all patients or right mTLE from HC, we found no such increase in accuracy.

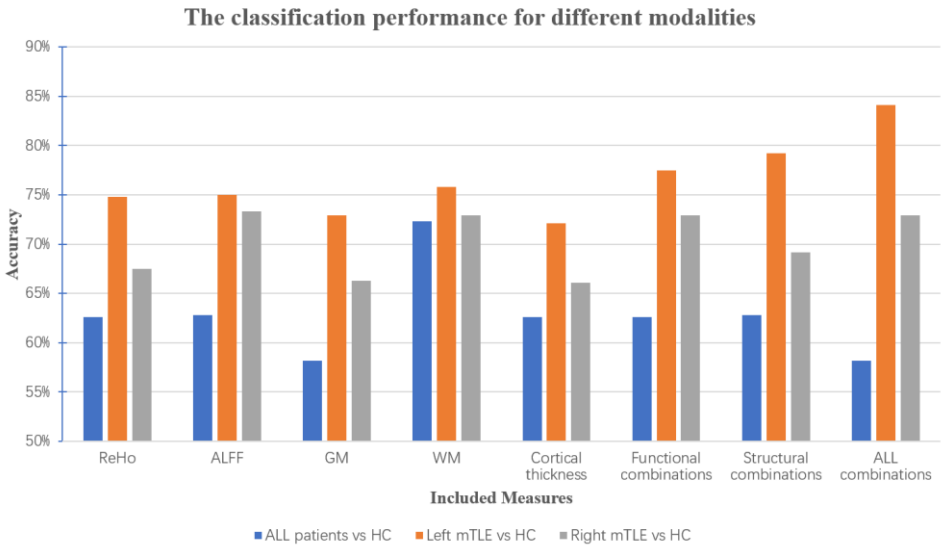

Figure 2. Overview of the classification accuracy based on different modalities. 'Functional combinations' means a multimodal SVM classifier using functional measures ReHo and ALFF as input; 'structural combinations' means a multimodal SVM classifier based on structural measures GM, WM and cortical thickness; 'ALL combinations' means a multimodal SVM classifier using all of the five measures ReHo, ALFF, GM, WM and cortical thickness as input.

Table 2. SVM classifier performance for the different modalities and combinations

	Accuracy	Sensitivity	Specificity	Recall	F_1 Score	AUC	P value
All patients vs HC							
ReHo	62.6%	67.5%	57.7%	61.5%	62.0%	64.1%	0.019
ALFF	62.8%	69.2%	56.4%	61.3%	62.0%	65.7%	0.017
GM	58.2%	56.6%	59.8%	58.5%	58.3%	61.4%	0.023
WM	72.3%	77.8%	66.8%	70.1%	71.2%	75.2%	0.012
Cortical thickness	62.6%	57.5%	67.7%	64.0%	63.3%	65.6%	0.026
Functional combinations [a]	62.6%	68.7%	56.5%	61.2%	61.9%	64.3%	0.023
Structural combinations [b]	62.8%	65.4%	60.2%	62.2%	62.5%	66.7%	0.019
All combinations [c]	58.2%	67.5%	48.9%	57.0%	57.6%	64.3%	0.035
Left mTLE vs HC							
ReHo	74.8%	80.7%	69.0%	72.2%	73.5%	77.8%	0.003
ALFF	75.0%	81.7%	68.3%	72.0%	73.5%	79.5%	0.001
GM	72.9%	72.5%	73.3%	73.1%	73.0%	74.1%	0.002
WM	75.8%	77.8%	73.7%	74.7%	75.2%	81.2%	0.009
Cortical thickness	72.1%	70.5%	73.7%	72.8%	72.4%	73.6%	0.002
Functional combinations [a]	77.5%	84.1%	70.8%	74.2%	75.8%	81.4%	0.001
Structural combinations [b]	79.2%	81.7%	76.7%	77.8%	78.5%	83.6%	0.001
ALL combinations [c]	84.1%	86.5%	81.7%	82.5%	83.3%	87.8%	0.001
Right mTLE vs HC							
ReHo	67.5%	61.7%	73.3%	69.8%	68.6%	71.3%	0.005
ALFF	73.3%	67.5%	79.2%	76.4%	74.8%	75.2%	0.002
GM	66.3%	55.0%	77.5%	71.0%	68.6%	68.2%	0.021
WM	72.9%	72.5%	73.3%	73.1%	73.0%	75.1%	0.002
Cortical thickness	66.1%	57.1%	75.1%	69.6%	67.8%	71.3%	0.017
Functional combinations [a]	72.9%	67.5%	78.3%	75.7%	74.3%	73.6%	0.002
Structural combinations [b]	69.2%	65.0%	73.3%	70.9%	70.0%	71.4%	0.006
All combinations [c]	72.9%	77.5%	68.3%	71.0%	71.9%	74.6%	0.002

[a] Functional combinations: multi-modal SVM classifier using functional measures ReHo and ALFF as input;
[b] Structural combinations: multi-modal SVM classifier based on structural measures GM, WM and cortical thickness;

[c] All combinations: multi-modal SVM classifier using all of the five measures ReHo, ALFF, GM, WM and cortical thickness as input.

Abbreviations: ReHo, Regional Homogeneity; ALFF, Amplitude Low Frequency Fluctuation; GM, Gray Matter; WM, White Matter; SVM, Support Vector Machine; mTLE, mesial Temporal Lobe Epilepsy; HC, Healthy Controls. AUC, Area Under the receiver operating characteristic Curve.

The most discriminative brain regions

In order to explore which brain regions contributed to single-subject classification, we computed the mean absolute values of the weights of the model across the different stages of the cross-validation, and then used a template mask based on the AAL atlas to extract the weight for each region. The top 10 brain regions with the highest mean values based on each measure are reported in Table 3.

The brain regions contributing to single-subject classification varied across our four measures of interest. However, some regions were detected at least in two of our four measures of interest (Figure 3 and Figure 4). In the classification of left mTLE and HC, regions which were detected in at least two individual measures included some structures of the left temporal lobe (such as left inferior temporal gyrus, left temporal pole, middle temporal gyrus) and of the DMN (default-mode network) [36, 37] (such as the left superior parietal gyrus, left inferior parietal and left angular gyrus). In the classification of right mTLE and HC, the main role was played by the right temporal lobe (right inferior temporal gyrus, right temporal pole: middle temporal gyrus, right temporal pole: superior temporal gyrus), left pallidum and bilateral putamen. Taken collectively, the discriminative regions for left mTLE and right mTLE mainly focused on the ipsilateral temporal lobe, and on extra-temporal regions such as DMN for left mTLE and left pallidum and bilateral putamen for right mTLE. When we used SVM to discriminate left mTLE and right mTLE from HC based on integration of the submaps of these regions for each measure, we obtained a balanced accuracy of 70.2% for left mTLE and 64.3% for right mTLE.

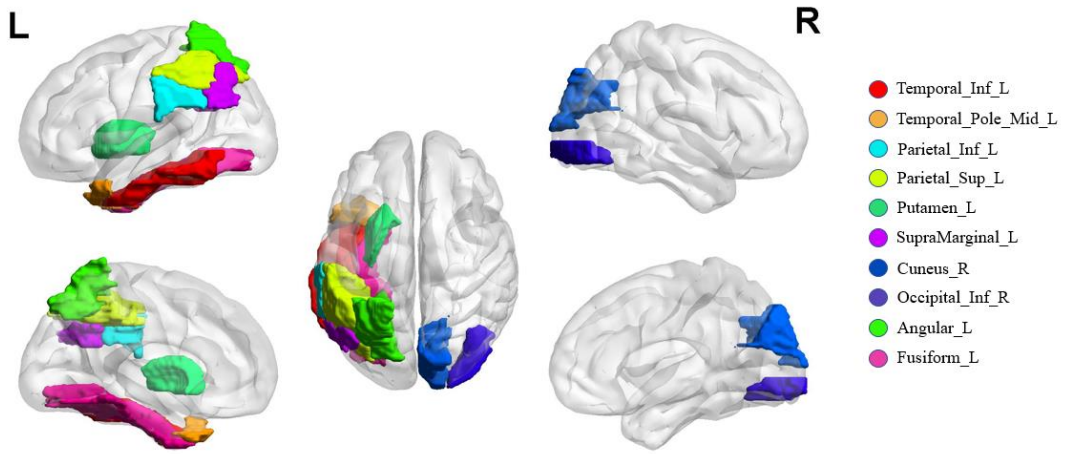

Figure 3. The distribution maps of the regions which were detected by at least two individual measures in the classification between left mTLE and healthy controls

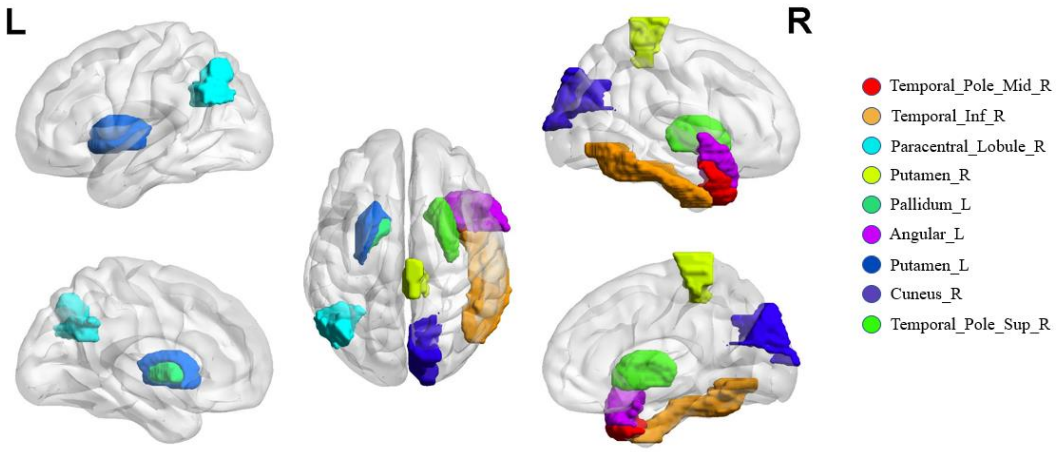

Figure 4. The distribution maps of the regions which were detected by at least two individual measures in the classification between right mTLE and healthy controls

Table 3. Ten brain regions making the greatest contribution to single-subject classification across the different measures.

Left mTLE VS HC	
ReHo	**Significance values**
Inferior temporal gyrus L	0.066
Pallidum L	0.064
Temporal pole: middle temporal gyrus L	0.058
Lingual gyrus R	0.053
Inferior occipital gyrus R	0.042
Superior parietal gyrus L	0.037
Inferior parietal gyrus L	0.036
Putamen L	0.035
Cuneus R	0.033
Supramarginal gyrus R	0.032
ALFF	**Significance values**
Superior parietal gyrus L	0.025
Precuneus L	0.021
Angular gyrus L	0.018
Inferior parietal gyrus L	0.018
Superior parietal gyrus R	0.017
Inferior occipital gyrus R	0.016
Supramarginal gyrus L	0.015
Cuneus R	0.015
Paracentral Lobule R	0.014
Postcentral gyrus R	0.014
GM	**Significance values**
Fusiform gyrus R	0.032
Angular gyrus L	0.028
Temporal pole: superior temporal gyrus L	0.022
Thalamus R	0.021
Putamen L	0.021
Cuneus L	0.019
Supramarginal gyrus L	0.019
Heschl's gyrus R	0.018
Cuneus R	0.015

Fusiform gyrus L	0.015
WM	**Significance values**
Temporal pole: middle temporal gyrus L	0.043
Parahippocampal gyrus L	0.035
Inferior temporal gyrus L	0.033
Paracentral lobule L	0.028
Superior temporal gyrus L	0.027
Superior occipital gyrus L	0.026
Fusiform gyrus L	0.021
Temporal pole: middle temporal gyrus R	0.021
Inferior occipital gyrus R	0.019
Parahippocampal gyrus R	0.018
Right mTLE VS HC	
ReHo	**Significance values**
Caudate L	0.078
Inferior temporal gyrus L	0.062
Temporal pole: middle temporal gyrus R	0.055
Inferior temporal gyrus R	0.053
Inferior occipital gyrus R	0.042
Paracentral lobule R	0.035
Putamen R	0.034
Pallidum L	0.031
Temporal pole: superior temporal gyrus L	0.030
Thalamus R	0.030
ALFF	**Significance values**
Precuneus L	0.035
Superior parietal gyrus L	0.029
Angular gyrus L	0.028
Inferior parietal gyrus L	0.026
Middle frontal gyrus (orbital part) R	0.025
Paracentral lobule R	0.021
Supramarginal gyrus L	0.020
Temporal pole: middle temporal gyrus R	0.020
Rectus R	0.019
Superior parietal gyrus R	0.019

GM	**Significance values**
Heschl gyrus R	0.028
Putamen L	0.027
Cuneus R	0.027
Temporal pole: superior temporal gyrus R	0.025
Fusiform gyrus R	0.023
Putamen R	0.022
Fusiform gyrus L	0.021
Pallidum L	0.021
Paracentral lobule R	0.020
Insula R	0.017
WM	**Significance values**
Inferior temporal gyrus R	0.036
Parahippocampal gyrus R	0.033
Temporal pole: middle temporal gyrus L	0.031
Middle temporal gyrus R	0.030
Temporal pole: superior temporal gyrus R	0.027
Putamen L	0.026
Putamen R	0.021
Cuneus R	0.020
Angular gyrus L	0.020
Parahippocampal gyrus L	0.020

All brain regions are identified using AAL (automated anatomical labeling); the vectors are computed using a template mask based on the AAL atlas to extract the absolute value of weight for each brain regions across the different folds of the cross-validation.

Abbreviations: ReHo, Regional Homogeneity; ALFF, Amplitude of Low Frequency Fluctuation; GM, Gray Matter; WM, White Matter; mTLE, mesial Temporal Lobe Epilepsy; HC, Healthy Controls; L, Left; R, Right.

Discussion

This study combined functional and structural MRI measures to distinguish mTLE patients from HC. Our results suggest that classification accuracy can be improved by dividing the mTLE patients into two groups (left and right) and by combining functional and structural MRI measures. The temporal lobe contributed most to the single subject classification, and some extra-temporal regions also had high discriminative power.

Consistent with our hypothesis (i), dividing the mTLE patients into left and right mTLE improved the classifier performance. This reflects the fact that left and right mTLE are associated with distinct brain alterations, corroborated by both functional and structural imaging findings [15-17]. Moreover, the accuracy of left mTLE versus HC was higher than right versus HC for different modalities, in accordance with previous ML studies [9, 14, 38]. One possible explanation, for which there is some evidence, is that the functional and structural alterations in left mTLE are more extensive than in right mTLE [39, 40]. In functional studies, left mTLE showed greater reduction of functional connectivity than right mTLE [41]. Furthermore, left mTLE has been reported to be associated with alteration of bilateral mesial temporal lobes, while right mTLE only with alteration of right mesial temporal lobe [42]. Some structural studies have found that left mTLE shows not only more extensive losses in white matter, but also more aberrant inter-tract correlations than right mTLE [43], while another study reports greater alteration of GM and WM in left mTLE [44]. The left hemisphere is of course dominant in most right-handed persons [45], and all of our subjects were right-handed. Thus the diversity may be explained on the basis that seizures originating in the dominant hemisphere cause more excitotoxic damage in left-hemisphere dominant patients.

Consistent with our hypothesis (ii), in the identification of left mTLE and HC, by combining two functional measures (ReHo, ALFF), we achieved an accuracy of 78% (comparable to 83% accuracy in distinguishing mTLE from HC reported in a study combining 6 rs-fMRI measures [12]). By combining 3 structural measures (GM, WM and cortical thickness) we achieved an accuracy of 79% (comparable to 81% accuracy in distinguishing mTLE from HC reported in a study combining multiparameter sMRI data) [11]. When we combined all the measures, we obtained the highest accuracy of 84%. This is accordance with the results reported in some other neuropsychiatric disorder [46, 47]. The increased accuracy supports the view that mTLE can cause not only structural abnormality but also functional alterations of brain [48-51]. Therefore, combining multi-modal measures within a single model appears to be a promising tool for improving classification of individual patients with mTLE. By contrast, in the classification of right mTLE and HC, the accuracy was not increase by combining functional and structural measures. The reason may be that some of the neuroimaging modalities (ReHo, GM, cortical thickness) we used were more sensitive in detecting left mTLE than right mTLE (Table 2).

The best-discriminative regions were widespread and not restricted to particular brain hemispheres or lobes across the four measures. There are two possible reasons in SVM why an individual region might display high discriminative power: a between-group feature value difference in that region; or a between-group difference in the correlation between that region and other areas. Thus, the widespread network revealed in this kind of study should not be interpreted in terms of individual regions, but as a spatially distributed pattern of a discrimination informed by all brain voxels. Direct comparison is difficult with reported sMRI or rs-fMRI studies using mass-univariate analyses, but it seems reasonable that brain regions showing greater

difference should contribute more in the SVM based classification. The discriminative regions we detected in more than two measures partially overlap with previous studies. For example, consistent with our hypothesis (iii), the ipsilateral temporal lobe contributed most to classification across the four measures. Consistent with this, in previous studies the epileptogenic zone often involved the mesial and lateral temporal lobe in mTLE [52, 53]. In the identification of left mTLE and HC, some regions of the DMN are also important. Previous studies have found functional or structural alterations of DMN in mTLE [19, 54, 55]. The DMN is an integrated system for self-related cognitive activity, including autobiographical, self-monitoring and social functions [56], so impairment of the DMN in mTLE may underlie the pathophysiological mechanism of impaired cognition [57]. Previous studies have suggested that alterations of DMN in mTLE may be related to the rich connections that exist between the hippocampus and several key structures of this network [58]. In addition, several subcortical regions, such as pallidum and putamen, also had high discriminative power in the classification of right mTLE and HC. Consistent with this, alterations of pallidum and putamen in mTLE have been reported in previous studies [59].

This study has some limitations. Firstly, a major challenge in the application of machine learning to high dimensional neuroimaging data is the risk of overfitting. We minimized this risk by using region-level features, which are associated with less noise and lower risk of overfitting, rather than voxel-level data [60]. Secondly, although the present results are promising, the development of a practical diagnostic will require several advances. The model will need even better accuracy, perhaps by including more diverse observations from multimodal imaging. Finally, to make it easier to discuss the neurobiology of mTLE, we identified the discriminative regions

based on the AAL atlas, with the potential drawback that some atlas areas (e.g. the hippocampus region) might be too large or unspecific to detect group differences.

In conclusion, the present study shows that dividing the mTLE patients into left and right mTLE, and combining multi-modal measures within a single model, both improve the classifier performance. We therefore suggest that subtyping of patients and integration of multi-modal neuroimaging modalities will be promising methods for improving classifier performance in the classification of individual patients with mTLE and HC.

Acknowledgements

This study was supported by the National Natural Science Foundation of China (Grant Nos. 81501452, 81621003, 81761128023, 81220108031 and 81227002), the Program for Chang jiang Scholars and Innovative Research Team in University (PCSIRT, grant IRT16R52) of China, the Chang jiang Scholar Professorship Award (Award No. T2014190) of China, and the CMB Distinguished Professorship Award (Award No. F510000/G16916411) administered by the Institute of International Education. And Du Lei was supported by the Newton International Fellowship Alumni Award from the Royal Society. The authors would like to thank all of the study participants and their families.

Compliance with ethics guideline

The authors declare that they have no conflicts of interest. This study was approved by the West China Hospital Clinical Trials and Biomedical Ethics Committee of

Sichuan University, and written informed consent was obtained from all of the participants. The study protocol was performed in accordance with the approved guidelines.

References

1. Engel J, Jr. Mesial temporal lobe epilepsy: what have we learned? Neuroscientist 2001;7(4):340-352.

2. Burianova H, Faizo NL, Gray M, Hocking J, Galloway G, Reutens D. Altered functional connectivity in mesial temporal lobe epilepsy. Epilepsy Res 2017;137:45-52.

3. Chen SG, Chen L, Huang HP, Lin WH. Relationship between resting state functional magnetic resonance imaging and memory function in mesial temporal lobe epilepsy. Neurol Sci 2017;372:117-125.

4. Memarian N, Thompson PM, Engel J, Jr., Staba RJ. Quantitative analysis of structural neuroimaging of mesial temporal lobe epilepsy. Imaging Med 2013;5(3):10-14

5. Liu M, Chen Z, Beaulieu C, Gross DW. Disrupted anatomic white matter network in left mesial temporal lobe epilepsy. Epilepsia 2014;55(5):674-682.

6. Kassahun Y, Perrone R, De Momi E, Berghofer E, Tassi L, Canevini MP, Spreafico R, Ferrigno G, Kirchner F. Automatic classification of epilepsy types using ontology-based and genetics-based machine learning. Artif Intell Med 2014;61(2):79-88.

7. Acharya UR, Yanti R, Wei ZJ, Krishnan MMR, Hong TJ, Martis RJ, Min LC. Automated Diagnosis Of Epilepsy Using Cwt, Hos And Texture Parameters. Int J Neural Syst 2013;23(3):1350009

8. Kerr WT, Cho AY, Anderson A, Douglas PK, Lau EP, Hwang ES, Raman KR, Trefler A, Cohen MS, Nguyen ST, Reddy NM, Silverman DH. Balancing Clinical and Pathologic Relevence in the Machine Learning Diagnosis of Epilepsy. Int Workshop Pattern Recognit Neuroimaging 2013;2013:86-89.

9. Focke NK, Yogarajah M, Symms MR, Gruber O, Paulus W, Duncan JS. Automated MR image classification in temporal lobe epilepsy. Neuroimage 2012;59(1):356-362.

10. Memarian N, Kim S, Dewar S, Engel J, Jr., Staba RJ. Multimodal data and

machine learning for surgery outcome prediction in complicated cases of mesial temporal lobe epilepsy. Comput Biol Med 2015;64:67-78.

11. Rudie JD, Colby JB, Salamon N. Machine learning classification of mesial temporal sclerosis in epilepsy patients. Epilepsy Res 2015;117:63-69.

12. Yang Z, Choupan J, Reutens D, Hocking J. Lateralization of Temporal Lobe Epilepsy Based on Resting-State Functional Magnetic Resonance Imaging and Machine Learning. Front Neurol 2015;6:184.

13. Del Gaizo J, Mofrad N, Jensen JH, Clark D, Glenn R, Helpern J, Bonilha L. Using machine learning to classify temporal lobe epilepsy based on diffusion MRI. Brain Behav 2017;7(10):e00801.

14. An J, Fang P, Wang W, Liu Z, Hu D, Qiu S. Decreased white matter integrity in mesial temporal lobe epilepsy: a machine learning approach. Neuroreport 2014;25(10):788-794.

15. Pereira FR, Alessio A, Sercheli MS, Pedro T, Bilevicius E, Rondina JM, Ozelo HF, Castellano G, Covolan RJ, Damasceno BP, Cendes F. Asymmetrical hippocampal connectivity in mesial temporal lobe epilepsy: evidence from resting state fMRI. BMC Neurosci 2010;11:66.

16. Kemmotsu N, Girard HM, Bernhardt BC, Bonilha L, Lin JJ, Tecoma ES, Iragui VJ, Hagler DJ, Jr., Halgren E, McDonald CR. MRI analysis in temporal lobe epilepsy: cortical thinning and white matter disruptions are related to side of seizure onset. Epilepsia 2011;52(12):2257-2266.

17. Kucukboyaci NE, Girard HM, Hagler DJ, Jr., Kuperman J, Tecoma ES, Iragui VJ, Halgren E, McDonald CR. Role of frontotemporal fiber tract integrity in task-switching performance of healthy controls and patients with temporal lobe epilepsy. J Int Neuropsychol Soc 2012;18(1):57-67.

18. Coan AC, Campos BM, Beltramini GC, Yasuda CL, Covolan RJ, Cendes F. Distinct functional and structural MRI abnormalities in mesial temporal lobe epilepsy with and without hippocampal sclerosis. Epilepsia 2014;55(8):1187-1196.

19. Liao W, Zhang Z, Pan Z, Mantini D, Ding J, Duan X, Luo C, Wang Z, Tan Q, Lu G, Chen H. Default mode network abnormalities in mesial temporal lobe

epilepsy: a study combining fMRI and DTI. Hum Brain Mapp 2011;32(6):883-895.

20. Wee CY, Yap PT, Zhang DQ, Denny K, Browndyke JN, Potter GG, Welsh-Bohmer KA, Wang LH, Shen DG. Identification of MCI individuals using structural and functional connectivity networks. Neuroimage 2012;59(3):2045-2056.

21. Scanlon C, Mueller SG, Cheong I, Hartig M, Weiner MW, Laxer KD. Grey and white matter abnormalities in temporal lobe epilepsy with and without mesial temporal sclerosis. J Neurol 2013;260(9):2320-2329.

22. Braga B, Yasuda CL, Cendes F. White matter atrophy in patients with mesial temporal lobe epilepsy: voxel-based morphometry analysis of T1- and T2-Weighted MR Images. Radiol Res Pract 2012;2012:481378.

23. Labate A, Cerasa A, Aguglia U, Mumoli L, Quattrone A, Gambardella A. Neocortical thinning in "benign" mesial temporal lobe epilepsy. Epilepsia 2011;52(4):712-717.

24. Zhang Z, Lu G, Zhong Y, Tan Q, Chen H, Liao W, Tian L, Li Z, Shi J, Liu Y. fMRI study of mesial temporal lobe epilepsy using amplitude of low-frequency fluctuation analysis. Hum Brain Mapp 2010;31(12):1851-1861.

25. Zhong J, Chen S, Ouyang Q, An D, Lu S. [Study on the ReHo in treatment-naive of temporal lobe epilepsy patients with depressive symptoms using resting functional-MRI (fMRI)]. Sheng Wu Yi Xue Gong Cheng Xue Za Zhi 2012;29(2):229-232.

26. Zeng H, Pizarro R, Nair VA, La C, Prabhakaran V. Alterations in regional homogeneity of resting-state brain activity in mesial temporal lobe epilepsy. Epilepsia 2013;54(4):658-666.

27. Proposal for revised classification of epilepsies and epileptic syndromes. Commission on Classification and Terminology of the International League Against Epilepsy. Epilepsia 1989;30(4):389-399.

28. Shorvon SD. The etiologic classification of epilepsy. Epilepsia 2011;52(6):1052-1057.

29. Fischl B, van der Kouwe A, Destrieux C, Halgren E, Segonne F, Salat DH,

Busa E, Seidman LJ, Goldstein J, Kennedy D, Caviness V, Makris N, Rosen B, Dale AM. Automatically parcellating the human cerebral cortex. Cerebral Cortex. 2004;14(1):11-22.

30. Desikan RS, Segonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, Albert MS, Killiany RJ. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 2006;31(3):968-980.

31. Cortes C, Vapnik V. Support-Vector Networks. Machine Learning 1995;20(3):273-297.

32. Chang CC, Lin CJ. LIBSVM: A Library for Support Vector Machines. ACM Trans Intell Syst Technol 2011;2(3):115-121

33. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E. Scikit-learn: Machine Learning in Python. J Mach Learn Res 2011;12:2825-2830.

34. Pettersson-Yeo W, Benetti S, Marquand AF, Joules R, Catani M, Williams SCR, Allen P, McGuire P, Mechelli A. An empirical comparison of different approaches for combining multimodal neuroimaging data with support vector machine. Front Neurosci. 2014;8:189

35. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 2002;15(1):273-289.

36. Damoiseaux JS, Rombouts SA, Barkhof F, Scheltens P, Stam CJ, Smith SM, Beckmann CF. Consistent resting-state networks across healthy subjects. Proc Natl Acad Sci U S A 2006;103(37):13848-13853.

37. Buckner RL, Andrews-Hanna JR, Schacter DL. The brain's default network: anatomy, function, and relevance to disease. Ann N Y Acad Sci 2008;1124:1-38.

38. Cantor-Rivera D, Khan AR, Goubran M, Mirsattari SM, Peters TM. Detection of temporal lobe epilepsy using support vector machines in multi-parametric

quantitative MR imaging. Comput Med Imaging Graph. 2015;41:14-28.

39. Kemmotsu N, Kucukboyaci NE, Cheng CE, Girard HM, Tecoma ES, Iragui VJ, McDonald CR. Alterations in functional connectivity between the hippocampus and prefrontal cortex as a correlate of depressive symptoms in temporal lobe epilepsy. Epilepsy Behav 2013;29(3):552-559.

40. McDonald CR, Hagler DJ, Ahmadi ME, Tecoma E, Iragui V, Gharapetian L, Dale AM, Halgren E. Regional neocortical thinning in mesial temporal lobe epilepsy. Epilepsia 2008;49(5):794-803.

41. Kucukboyaci NE, Kemmotsu N, Cheng CE, Girard HM, Tecoma ES, Iragui VJ, McDonald CR. Functional connectivity of the hippocampus in temporal lobe epilepsy: feasibility of a task-regressed seed-based approach. Brain Connect. 2013;3(5):464-474.

42. Ji GJ, Zhang ZQ, Zhang H, Wang J, Liu DQ, Zang YF, Liao W, Lu GM. Disrupted Causal Connectivity in Mesial Temporal Lobe Epilepsy. Plos One. 2013;8(5): e63183

43. Pustina D, Doucet G, Sperling M, Sharan A, Tracy J. Increased microstructural white matter correlations in left, but not right, temporal lobe epilepsy. Hum Brain Mapp. 2015;36(1):85-98.

44. Lu JJ, Li WJ, He HG, Feng F, Jin ZY, Wu LW. Altered hemispheric symmetry found in left-sided mesial temporal lobe epilepsy with hippocampal sclerosis (MTLE/HS) but not found in right-sided MTLE/HS. Magn Reson Imaging. 2013;31(1):53-59.

45. Lishman WA, McMeekan ER. Handedness in relation to direction and degree of cerebral dominance for language. *Cortex.* 1977;13(1):30-43.

46. Zhang Q, Wu Q, Zhu H, He L, Huang H, Zhang J, Zhang W. Multimodal MRI-Based Classification of Trauma Survivors with and without Post-Traumatic Stress Disorder. Front Neurosci. 2016;10:292.

47. Hinrichs C, Singh V, Xu G, Johnson SC, Alzheimers Disease Neuroimaging I. Predictive markers for AD in a multi-modality framework: an analysis of MCI progression in the ADNI population. Neuroimage. 2011;55(2):574-589.

48. Barron DS, Tandon N, Lancaster JL, Fox PT. Thalamic structural connectivity

in medial temporal lobe epilepsy. Epilepsia. 2014;55(6):e50-55.

49. Peng BW, Wu LW, Zhang LH, Chen Y. The relationship between hippocampal volumes and nonverbal memory in patients with medial temporal lobe epilepsy. Epilepsy Res. 2014;108(10):1839-1844.

50. Sequeira KM, Tabesh A, Sainju RK, DeSantis SM, Naselaris T, Joseph JE, Ahlman MA, Spicer KM, Glazier SS, Edwards JC, Bonilha L. Perfusion Network Shift during Seizures in Medial Temporal Lobe Epilepsy. Plos One. 2013;8(1):e53204.

51. Bertram EH. Extratemporal lobe circuits in temporal lobe epilepsy. Epilepsy Behav. 2014;38:13-18.

52. Sankar T, Bernasconi N, Kim H, Bernasconi A. Temporal lobe epilepsy: differential pattern of damage in temporopolar cortex and white matter. Hum Brain Mapp. 2008;29(8):931-944.

53. Curia G, Lucchi C, Vinet J, Gualtieri F, Marinelli C, Torsello A, Costantino L, Biagini G. Pathophysiogenesis of Mesial Temporal Lobe Epilepsy: Is Prevention of Damage Antiepileptogenic? Curr Med Chem 2014;21(6):663-688.

54. Liao W, Zhang Z, Pan Z, Mantini D, Ding J, Duan X, Luo C, Lu G, Chen H. Altered functional connectivity and small-world in mesial temporal lobe epilepsy. PLoS One 2010;5(1):e8525.

55. Pittau F, Grova C, Moeller F, Dubeau F, Gotman J. Patterns of altered functional connectivity in mesial temporal lobe epilepsy. Epilepsia 2012;53(6):1013-1023.

56. Menon V. Large-scale brain networks and psychopathology: a unifying triple network model. Trends Cogn Sci. 2011;15(10):483-506.

57. Zhang Z, Lu G, Zhong Y, Tan Q, Liao W, Wang Z, Wang Z, Li K, Chen H, Liu Y. Altered spontaneous neuronal activity of the default-mode network in mesial temporal lobe epilepsy. Brain Res. 2010;1323:152-160.

58. Cataldi M, Avoli M, de Villers-Sidani E. Resting state networks in temporal lobe epilepsy. Epilepsia. 2013;54(12):2048-2059.

59. Bouilleret V, Semah F, Chassoux F, Mantzaridez M, Biraben A, Trebossen R, Ribeiro MJ. Basal ganglia involvement in temporal lobe epilepsy: a functional and morphologic study. Neurology. 2008;70(3):177-184.

60. Vieira S, Pinaya WHL, Mechelli A. Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: Methods and applications. Neurosci Biobehav Rev. 2017;74:58-75.