Epilepsy Research 117 (2015) 63–69

Contents lists available at www.sciencedirect.com

Epilepsy Research

journal homepage: www.elsevier.com/locate/epilepsyres

Machine learning classification of mesial temporal sclerosis in epilepsy patients

Jeffrey D. Rudie[a], John B. Colby[a], Noriko Salamon[a,b,*]

[a] David Geffen School of Medicine at UCLA, United States
[b] Department of Radiology, Ronald Reagan Hospital, UCLA, United States

ARTICLE INFO

Article history:
Received 13 May 2015
Received in revised form 13 August 2015
Accepted 7 September 2015
Available online 9 September 2015

Keywords:
Machine learning
Mesial temporal sclerosis
Cortical thickness
Hippocampus
Cortical folding
Temporal lobe epilepsy

ABSTRACT

Background and purpose: Novel approaches applying machine-learning methods to neuroimaging data seek to develop individualized measures that will aid in the diagnosis and treatment of brain-based disorders such as temporal lobe epilepsy (TLE). Using a large cohort of epilepsy patients with and without mesial temporal sclerosis (MTS), we sought to automatically classify MTS using measures of cortical morphology, and to further relate classification probabilities to measures of disease burden.
Materials and methods: Our sample consisted of high-resolution T1 structural scans of 169 adults with epilepsy collected across five different 1.5 T and four different 3 T scanners at UCLA. We applied a multiple support vector machine recursive feature elimination algorithm to morphological measures generated from FreeSurfer's automated segmentation and parcellation in order to classify Epilepsy patients with MTS ($n = 85$) from those without MTS ($N = 84$).
Results: In addition to hippocampal volume, we found that alterations in cortical thickness, surface area, volume and curvature in inferior frontal and anterior and inferior temporal regions contributed to a classification accuracy of up to 81% ($p = 1.3 \times 10^{-17}$) in identifying MTS. We also found that MTS classification probabilities were associated with a longer duration of disease for epilepsy patients both with and without MTS.
Conclusions: In addition to implicating extra-hippocampal involvement of MTS, these findings shed further light on the pathogenesis of TLE and may ultimately assist in the development of automated tools that incorporate multiple neuroimaging measures to assist clinicians in detecting more subtle cases of TLE and MTS.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

During the diagnostic workup of epilepsy, the identification of major structural abnormalities, including mesial temporal sclerosis (MTS), is critical for guiding clinical decision-making. In the past two decades, more advanced imaging acquisition and analysis methods have been used to detect more subtle morphological abnormalities in epilepsy patients. However, a majority of these studies identify a group-level difference, which has little clinical utility. Therefore, diagnosis of MTS and other structural abnormalities continues to be based on visual inspection by trained neuroradiologists. More advanced machine-based learning methods can be applied to individual subjects and have significant potential for assisting with diagnosis and predicting treatment response in individuals with epilepsy and other neurological illnesses.

The majority of neuroimaging research has focused on temporal lobe epilepsy (TLE), which is the most prevalent form of medically intractable epilepsy (Engel, 1996). The pathologic finding of MTS exists in up to 65% of cases of TLE (Babb et al., 1984). MTS is characterized histologically by cellular loss and hippocampal reorganization and is often identified on MRI by hippocampal atrophy and signal abnormalities (Berkovic et al., 1991).

A multitude of prior neuroimaging studies have examined whole-brain differences in cortical morphology between patients with TLE and controls using both voxel based morphometry (VBM) and cortical thickness methods. As reviewed by Keller and Roberts (2008), VBM studies have found the largest effects in nearby ipsilateral medial temporal cortex, as well as more widespread effects in regions including the thalamus and frontal and parietal lobes. The

Abbreviations: TLE, temporal lobe epilepsy; MTS, mesial temporal sclerosis; VBM, voxel based morphometry; SVM, support vector machine; RFE, recursive feature elimination; ROC, receiver operator curve.

* Corresponding author at: 757 Westwood Blvd, Suite 1621D, Los Angeles, CA 90095, United States.
E-mail address: nsalamon@mednet.ucla.edu (N. Salamon).

http://dx.doi.org/10.1016/j.eplepsyres.2015.09.005
0920-1211/© 2015 Elsevier B.V. All rights reserved.

findings of more widespread alterations in cortical morphology that extend beyond the hippocampus are consistent with TLE's comorbid deficits in executive, intellectual and language functioning (Oyegbile et al., 2004). More recently, using methods that reconstruct the cortical surface to more precisely measure gray matter structure, researchers found up to 30% reductions in cortical thickness bilaterally in multiple frontal, temporal and occipital regions in TLE patients with MTS compared to controls (Lin et al., 2007; Bernhardt et al., 2010; Kemmotsu et al., 2011). These findings are present in TLE patients with and without MTS but have been shown to be stronger in patients with MTS (Labate et al., 2011). Additionally, these methods allow for measurements of cortical folding and complexity, which have also been reported to be abnormal in individuals with TLE and MTS, particularly in ipsilateral temporal and frontal cortices (Voets et al., 2011; Alhusaini et al., 2012).

These computer-based methods have improved our understanding of the neurobiology of TLE. However, relative to the classic findings of hippocampal atrophy and T2 signal abnormalities, the alterations are more subtle, distributed and variable across patients, which makes visual assessment by neuroradiologists a more challenging task. Additionally, in order for the methods to have clinical utility they need to be able to be applied to individual subjects, yet these approaches generally report differences at the group level. Thus, an automated tool that incorporates all of these findings into a single metric that predicts MTS or TLE has more promise in supplementing the visual analysis of neuroradiologists. Machine learning approaches seek to incorporate multiple data points in order to build a model that can make predictions in future data sets. Support vector machines (SVM) are a class of machine learning algorithms that are well suited for neuroimaging data as they are fast, flexible and can be readily automated. They have been studied extensively in computer science and have been applied to neuroimaging data from a variety of neurological diseases including Alzheimer's disease (Klöppel et al., 2008), autism spectrum disorder (Ecker et al., 2010) and ADHD (Colby et al., 2012).

A few prior investigations have applied machine-learning methods to TLE. Using a linear discriminant analysis McDonald et al. (2008) showed that lateral temporal cortical thickness could significantly discriminate TLE patients from controls with 74% accuracy. By focusing solely on quantification of hippocampal volume and T2 signal changes, computer-aided methods were shown to improve detection of MTS by 28% compared to visual analysis alone (Coan et al., 2014). Using a voxel-based approach for white and gray matter segmentation as well as voxel based DTI measures, Focke et al. (2012) were able to distinguish TLE with MTS from controls with greater than 88% accuracy. Cantor-Rivera et al. (2015) used multiple measures, including quantitative regional T1/T2 and DTI values to classify TLE patients from controls with similarly high classification accuracies. Another recent study (Bernhardt et al., 2015) used a more advanced classification approach that identified multiple TLE subtypes showing distinct patterns of structural abnormalities that were further related to surgical outcomes.

In this study, we applied an SVM algorithm to measures of brain morphology, including cortical thickness, volume, and curvature, generated from FreeSurfer's automated segmentation and parcellation to a large sample of epilepsy patients at UCLA in order to identify MTS. Additionally, we related MTS classification scores with clinical measures, including age of onset, disease duration and seizure frequency in order to glean further insights into the pathogenesis and natural course of TLE and MTS.

2. Materials and methods

2.1. Subjects

Our sample consisted of high-resolution T1 structural scans of 169 adults scanned under UCLA's epilepsy protocol (IRB # 11-001678). There were a total of 848 individuals scanned across five different 1.5 T scanners (Avanto, Signa Genesis, Signa HD, two Sonata) and four different 3 T scanners (Verio, Skyra and two Trio-Tim) under UCLA's epilepsy protocol between April 2003 and May 2013. In order to be included in our study, individuals had to have a clinical diagnosis of epilepsy, either with or without MTS, and without evidence of other major structural abnormalities including isolated cortical dysplasia, tuberous sclerosis, leukomalacia, glioma, gray matter heterotopia or encephalomalacia as verified by a board certified neuroradiologist (N.S.). There were a total of 84 patients that had a clinical diagnosis of epilepsy without evidence of any structural abnormalities and which served as the control group. There were a total of 85 patients that had a diagnosis of epilepsy with left, right or bilateral MTS without any other structural abnormalities. Of the patients with MTS, 42 had left MTS, 35 had right MTS and 8 had bilateral MTS (Three of which were described to have left greater than right MTS). There were no significant differences ($p > 0.05$) between the control and MTS groups when comparing age, scanner strength (1.5 T vs. 3 T) and data resolution as measured by voxel volume (See Table 1 for group characteristics).

A retrospective chart review was done on all 169 patients in order to determine age of seizure onset, duration of disease and seizure frequency. In estimating seizure frequency, the average numbers of seizures were taken from chart reviews and/or a standardized clinical assessment written for each patient being evaluated by UCLA Adult Epilepsy Program. In cases where patients had multiple types of seizures, the total seizure frequency was summed across different seizure types. There were differences between groups for demographics including age of seizure onset, duration of disease, and seizure frequency, such that patients with MTS had an earlier onset of disease ($p = 0.001$) and disease duration ($p = 0.0001$) but lower seizure frequency ($p = 0.01$) than patients without MTS (Table 1).

2.2. MRI preprocessing

The basic T1-weighted anatomical MPRAGE sequence used across different scanners was as follows: TR/TE/TI = 1900/2.89/900 ms, 9° flip angle, 0.98 mm 0.98 mm 1.0–1.8 mm slice thickness. The exact parameters varied slightly across scanners. For additional details regarding acquisition parameters please see Lin et al. (2007), which used an overlapping dataset. The T1-weighted anatomical MRI scans were processed with FreeSurfer's *recon-all* processing pipeline for cortical reconstruction and volumetric segmentation (Fischl and Dale, 2000; Fischl, 2004) (software freely available at http://surfer.nmr.mgh.harvard.edu/). This method automatically generates reliable volume and thickness segmentations of white matter, gray matter, and subcortical volumes. The streamlined pipeline included removal of non-brain tissue, Tailarach transformations, segmentation of subcortical white and deep gray matter regions, intensity normalization and atlas registration. After these steps, 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 (Desikan et al., 2006). Importantly, it has been shown that the normalization process performed by FreeSurfer has good test–retest reliability across field strengths and scanner manufacturers (Han et al., 2006; Pfefferbaum et al., 2012).

Nine measures for each of the 34 cortical regions were calculated per hemisphere. These measures consisted of surface area, gray matter volume, average cortical thickness, cortical thickness standard deviation, cortical mean curvature, Gaussian curvature, cortical folding index, cortical curvature index and number of vertices. Additionally, three morphological measures (regional volume in mm^3, regional voxel intensity mean, and regional voxel intensity standard deviation) were calculated for 45 non-cortical regions.

Table 1
Mean and standard deviation of sample demographics.

Characteristic	Epilepsy with MTS	Epilepsy with normal MRI	*p* value
Sample Size	85	84	
Age (years)	32.9 ± 12.3	35.6 ± 13.3	0.16
Gender	44 male, 41 female	40 male, 44 female	0.59
1.5 T or 3 T Scanner	21 3 T, 64 1.5 T	23 3 T, 61 1.5 T	0.69
MPRAGE voxel volume (mm^3)	1.06 ± 0.59	1.09 ± 0.55	0.71
Age at seizure onset	11.7 ± 12.1	17.5 ± 12.5	0.001
Years since diagnosis	24.0 ± 13.0	15.5 ± 10.8	0.00001
Average seizures/month	13.9 ± 20.0	23.7 ± 29.6	0.01
Estimated life seizures	4147 ± 6877	4543 ± 7327	0.72

Data is mean ± standard deviation, minimum–maximum. Columns on the right display *p*-values for two sample *t*-tests *p*-values for each sample characteristic except for gender, which displays *p*-values from a Chi square test.

These regions included subcortical regions, white matter regions, ventricles, and other non-gray matter entities (i.e. white matter hyper-intensities). Of these non-cortical regions, 13 were midline structures and 16 were lateralized. Thus, there was total of 747 measures per subject overall (612 cortical and 135 non-cortical) and 354 lateralized measures per subject per hemisphere (306 cortical and 48 non-cortical) that were available for use as features in the machine learning analysis.

2.3. Machine learning methods and analysis

All machine-learning analyses were performed using R (http://www.r-project.org/) and the tools are freely available at http://github.com/johncolby/SVM-RFE. The methods were implemented in a similar fashion as Colby et al. (2012). We applied linear support vector machine recursive feature elimination (SVM-RFE; Guyon et al., 2002) in order to obtain a ranked list of features that best distinguished the 85 epilepsy patients with defined MTS from the 84 without. We included scanner strength, voxel size, patient age and gender as covariates in the analysis. The SVM-RFE method allows one to minimize redundant and extraneous features that could potentially degrade classifier performance (Farahat et al., 2011). SVM-RFE works backwards from the initial set of features and eliminates the least "useful" feature on each recursive pass. SVM-RFE has been applied successfully for feature selection across several functional neuroimaging studies (De Martino et al., 2008; Craddock et al., 2009).

To calculate the performance of the classifier, we used a 10-trial cross validation method, whereby the whole sample was randomly broken into 10 groups and a classifier was trained on 9/10th of the sample and then tested on the remaining 1/10th of the sample, across a total of ten trials. Within each of the ten trials, the classification accuracy on the holdout sample was gauged while varying the number of the top features used as input. Averaging across all 10 trials allowed us to generate plots of generalization performance vs. number of features. Importantly, this estimation of generalization performance and 10-trial cross validation was external to the feature ranking step. Thus, the features were ranked 10 times, each time independent of the hold-out samples for that given trial. This ensured that the estimated generalization performance is unbiased by spurious features that might explain the training class labels but might not generalize to the population (Ambroise and McLachlan, 2002).

The SVM-RFE was first applied to all 747 features from both hemispheres to discriminate the 85 Epilepsy patients with MTS from the 84 patients without MTS (Fig. 1A). The SVM-RFE was then applied using only the 354 features from the ipsilateral hemisphere with identified MTS relative to control patients (Fig. 1C). In patients with bilateral MTS, values from the left hemisphere were used given the fact that three of the bilateral cases were described to have findings of left greater than right MTS. In the ipsilateral analysis, left sided features were used from control subjects given that the majority of MTS cases were left sided. The true positive rate was plotted as a function of the false positive rate to generate receiver operator curves (ROC) for the whole brain and ipsilateral analyses (Fig. 1B and D). In order to more easily visualize and interpret the main findings, the top twenty of the 143 features useful features from the ipsilateral analysis are shown in Table 2 and Fig. 2 (all 142 features needed for peak classification in the ipsilateral analysis are included in Table S1).

Supplementary Table S1 related to this article can be found, in the online version, at http://dx.doi.org/10.1016/j.eplepsyres.2015.09.005.

The number of features that yielded the highest accuracy, based on the RFE, was used to generate classification probabilities (scores ranging from 0 to 1, with 1 representing 100% probability of MTS) for each subject averaged across the ten trials for the ipsilateral analysis. The classification scores were then correlated with age of onset of disease, disease duration, seizure frequency and lifetime seizure for all subjects and each group separately (Fig. 3). Partial correlations were also performed to control correlations with age of onset by disease duration and correlations with disease duration by age of onset. All statistical analyses performed on demographics and clinical measures with classification probabilities were conducted with PASW Statistics 18, Release Version 18.0.3 (SPSS, Inc., 2009, Chicago, IL).

3. Results

3.1. Classification accuracy

When using all 747 features from both hemispheres to distinguish patients with MTS from controls, we were able to achieve a

Table 2
Top 20 features used in classification.

Characteristic	Average ranking
Hippocampus volume	1.4
Pars orbitalis thickness	13.8
Frontal pole gray volume	14.2
STS bank surface area	16.9
Pars triangularis surface area	17.9
Entorhinal surface area	20.2
Medial orbital frontal folding index	25.1
Lingual gyrus surface area	26.7
Caudal anterior cingulate surface area	30.1
Posterior cingulate surface area	33.0
Pars orbitalis mean curvature	34.7
Cuneus Gaussian curvature	43.2
Inferior temporal gray volume	47.3
Parahippocampal thickness	49.9
Inferior temporal surface area	51.3
Insula curving index	51.3
Temporal pole curving index	52.5
Supramarginal gyrus thickness	54.7
Lingual gyrus gray volume	58.9

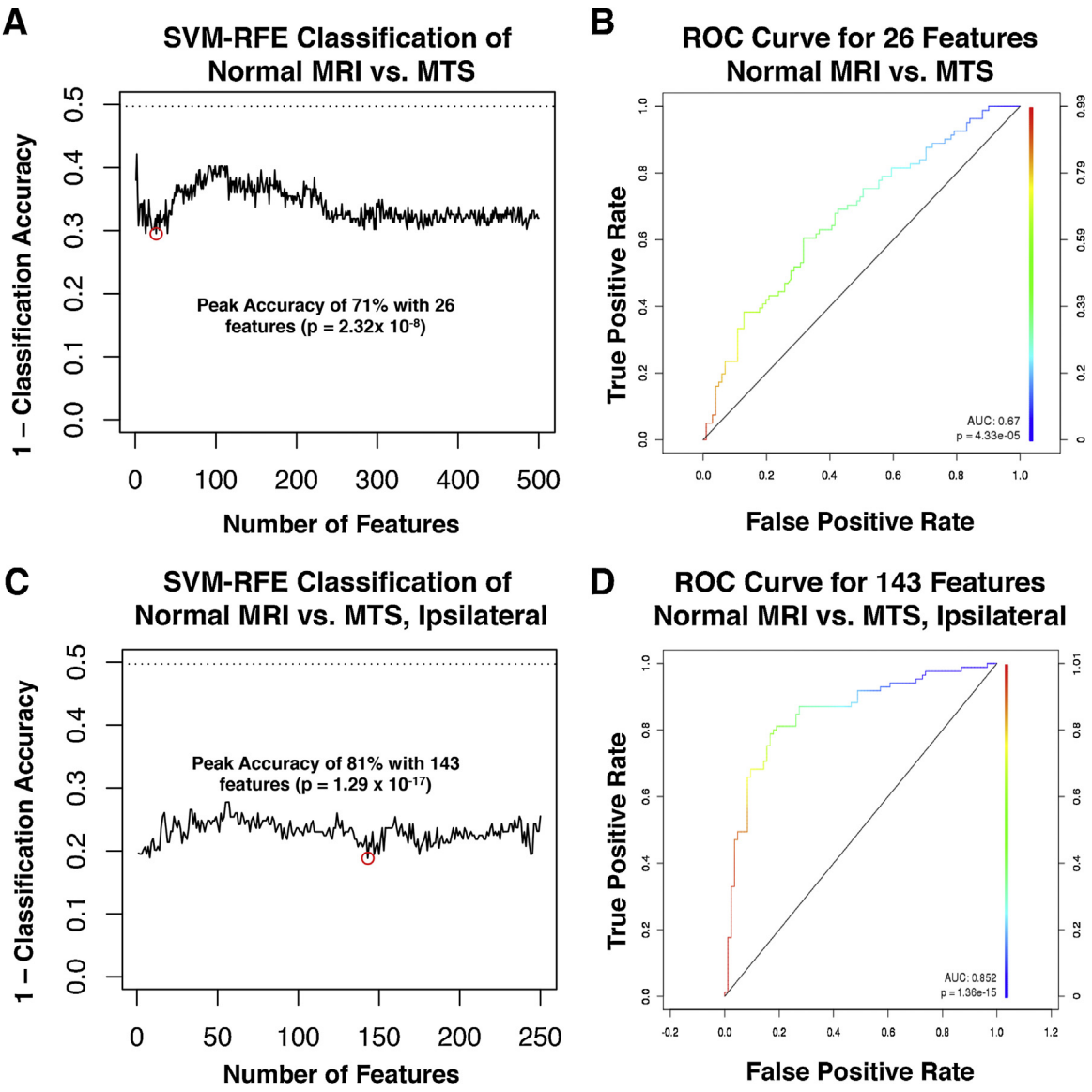

Fig. 1. Results of the support vector machine recursive feature elimination (SVM-RFE) analyses comparing epilepsy patients with mesial temporal sclerosis (MTS) to epilepsy patients with normal MRI's (control). SVM-RFE results for MTS vs. control for whole-brain (A) and ipsilateral hemisphere (C) analyses are displayed where 1-classification accuracy is plotted as function of the number of top features used. Receiver operator curves for whole brain (B) and ipsilateral hemisphere (D) analyses using the number of features that resulted in the highest accuracy are displayed with the true positive rate plotted as a function of the false positive rate.

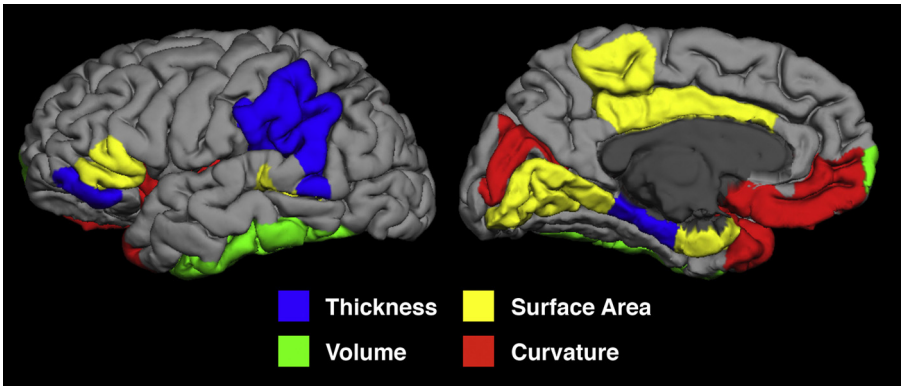

Fig. 2. Surface-based rendering of the top twenty morphological measures that distinguish epilepsy patients with MTS from epilepsy patients without MTS in the ipsilateral SVM-RFE analysis. The colors represent different types of cortical morphological features including cortical thickness (blue), surface area (yellow), volume (green) and curvature (red). Please see Table 2 for the names and rankings of each cortical region shown in this visualization.

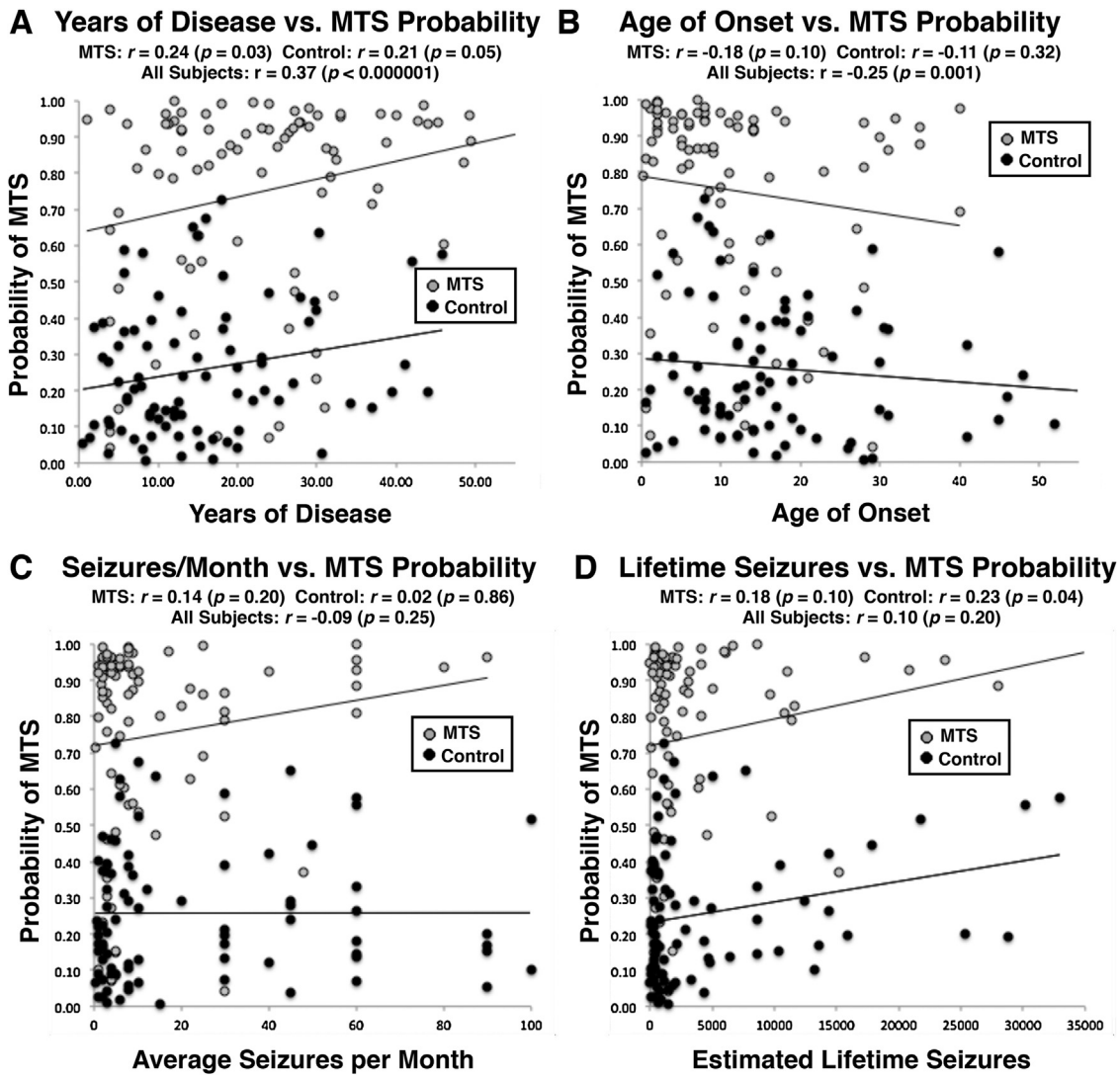

Fig. 3. Relationship between MTS classification probability and clinical measures. The probability of MTS is displayed as a function of years of disease (A), age of onset (B), seizures/month (C) and estimated lifetime seizures (D) for the MTS (gray) and control (black) groups.

peak classification accuracy of 71% with 26 features ($p = 2.32 \times 10^{-8}$; based on a binomial distribution coin toss; Fig. 1A). An ROC curve using probabilities from this analysis found an area under the curve of 0.67 ($p = 4.33 \times 10^{-5}$; Fig. 1B).

When comparing only the 354 features from the ipsilateral hemisphere with MTS, peak classification accuracy reached 81% with 143 features ($p = 1.29 \times 10^{-17}$; see Fig. 1C). An ROC curve using probabilities from this analysis found an area under the curve of 0.85 ($p = 1.36 \times 10^{-15}$; Fig. 1D).

3.2. Features contributing to classification accuracy

The top twenty of the 142 features used to obtain peak classification accuracy in the lateralized analysis are listed in Table 2 (with all 142 features listed in Table S1). The most reliable feature across the different trials was hippocampal volume. When performing a post hoc analysis of variance, we found that hippocampal volume explained 46% of the variance in MTS probability scores (Adjusted R squared = 0.464). Additional features useful for classification included cortical thickness, volume and surface area of nearby temporal and frontal regions. These top twenty features are displayed on the FreeSurfer surface-rendered brain with different colors representing different aspects of altered morphology overlaid on the corresponding brain regions (Fig. 2).

3.3. Correlations with clinical measures

A longer duration of disease was associated with a significantly higher probability of MTS across all subjects ($r = 0.37$, $p < 0.000001$, 2-tail), within the MTS subjects ($r = 0.24$, $p = 0.03$, 2-tail) and within the control group ($r = 0.21$, $p = 0.05$, 1-tail; Fig. 3A). These correlations remained significant when controlling for age of onset across all subjects ($r = 0.29$, $p = 0.0001$, 2-tail), and trends for this relationship remained within MTS subjects ($r = 0.20$, $p = 0.06$, 2-tail) and within the control group ($r = 0.18$, $p = 0.10$, 2-tail).

An earlier age of disease onset was associated with a significantly higher probability of MTS across all subjects ($r = -0.25$, $p = 0.001$, 2-tail) and there was a trend for this relationship within MTS subjects ($r = -0.18$, $p = 0.10$, 2-tail; Fig. 3B). When controlling for duration of disease, the correlation between age of onset and MTS probability was no longer significant.

There were no significant relationships identified between seizure frequency and MTS probability (Fig. 3C). The estimated amount of lifetime seizures was associated with a significantly higher probability of MTS within the control group ($r = -0.23$,

$p = 0.04$, 2-tail; Fig. 3D), with a trend for this relationship within the MTS subjects ($r = 0.18$, $p = 0.10$, 2-tail).

4. Discussion

The correct identification of structural brain abnormalities, including MTS, directly affects the clinical management of patients with intractable epilepsy. Our findings show significant promise in automatically detecting MTS using a machine learning approach on structural neuroimaging data. Despite heterogeneity introduced in our sample by including data from multiple different scanners, we were able to automatically identify MTS in epilepsy patients with up to 71% accuracy for whole-brain comparisons and up to 82% accuracy when focusing on the hemisphere with MTS. The distribution of morphological abnormalities that contributed to classification accuracy is consistent with prior literature and the tendency for an individual to be classified as having MTS was correlated with a longer duration of disease, providing further insight into the natural course of the disease.

The cortical effects of MTS have been reported to be more predominant ipsilateral to the side of MTS (Labate et al., 2011). As we found in this study, the classifier was more accurate and there was a greater area under the ROC curve for the analysis focusing only on the hemisphere ipsilateral to the side of MTS (Fig. 1). Additionally, the SVM-RFE algorithm reached peak accuracy using a larger number of features for the analysis focusing on the ipsilateral side. This suggests the possibility that including additional data with little signal (i.e., the hemisphere without MTS) may add noise and degrade the classifiers ability to identify MTS. Interestingly, Cantor-Rivera et al. (2015) found that subdividing their analysis into left and right TLE subjects increased their accuracy. Additionally, prior studies have found a differential distribution of left vs. right-sided MTS effects on cortical thickness (Lin et al., 2007). Therefore, a classification tool developed for clinical use would ideally generate classification scores for left and right hemispheres individually.

Consistent with the neuroradiologic diagnosis of MTS, the feature that contributed most to classification accuracy was hippocampal volume (Table 2). Other morphological characteristics that contributed to classification accuracy included altered thickness, surface area and curvature in inferior frontal and anterior and inferior temporal regions (Table 2, Fig. 2). Similar to previous studies on cortical thickness, the distribution of effects was relatively broad but had a predominance of effects in nearby frontal, temporal and limbic regions (Lin et al., 2007; Bernhardt et al., 2010; Kemmotsu et al., 2011). Additionally, measures of curvature and folding were found to be abnormal in ipsilateral temporal and frontal cortices, concordant with prior work (Voets et al., 2011; Alhusaini et al., 2012).

We also found that classification scores (probability of an MTS diagnosis) were correlated with clinical measures, including age of onset and duration of disease (Fig. 3). There was an association between MTS probability and earlier age of onset as well as a longer duration of disease. Given the correlation between these two clinical measures, we performed a partial correlation analysis, which showed that only the relationship between MTS probability and duration of disease remained significant when controlling for the other measure. In support of this finding, previous studies have found relationships between cortical thickness and duration of disease but not between cortical thickness and age of onset (Lin et al., 2007; McDonald et al., 2008). In contrast, Hermann et al., 2003 found that age of onset was a more important factor than duration disease in explaining differences in callosal volume in TLE patients.

The combination of a TLE diagnosis from clinical and EEG findings with the identification of MTS is a classic constellation of findings. However, it is unclear whether TLE without MTS is a distinct entity from TLE with MTS. Previous work has shown that TLE patients without MTS have a similar topology of morphological alterations including in temporal, medial frontal and cingulate regions as those without MTS (Bernhardt et al., 2010), but perhaps just to a lesser degree (Labate et al., 2011). We found that the MTS patients in our sample had an earlier onset and longer duration of disease, yet lower seizure frequency than the epilepsy patients without MTS. Additionally, even in the control patients that did not have MTS, a longer duration of disease was positively correlated with MTS classification probabilities. Taken together, these findings suggest the possibility that, on average, MTS may represent more of an end-stage of TLE, such that MTS may develop as the result of many seizures over multiple years. Albeit, there is a lot of heterogeneity within TLE and some patients are diagnosed with MTS only after a short duration of epilepsy. This leaves open the possibility that MTS could itself be pathological, causing increased seizures, at least in some cases. Additionally, MTS identification is still an important prognostic factor since it has been shown that medical treatment will be less effective and it is a positive indicator for outcomes after surgery (Téllez-Zenteno et al., 2010). Thus, regardless of whether TLE with MTS and without MTS are distinct entities, more accurately identifying MTS should assist with clinical decision-making.

Our investigation has several limitations and future directions that are worth noting. In contrast to most prior neuroimaging studies examining epilepsy patients, our study did not use healthy subjects as the control group; instead we utilized other epilepsy patients without MTS (or other structural abnormalities). Using a healthy control sample would likely have drastically improved overall accuracy by magnifying the differences between healthy and diseased brains. However, using healthy control subjects is less generalizable to the clinical scenario in which one must identify MTS in patients with known epilepsy in order to determine who among different epilepsy patients would most benefit from surgery vs. other types of treatments. Overall, the distribution of our findings is consistent with other studies and with the notion that there is a spectrum of differences between healthy control, TLE without MTS, and TLE with MTS. As demonstrated by Bernhardt et al. (2015) classification accuracy and correlations with surgical outcomes could be improved in future work by teasing apart heterogeneity in TLE samples by identifying subgroups based on imaging markers.

In this study, we sought to use the largest sample possible by including all data collected at UCLA in a given time period. Therefore, we included neuroimaging data across an array of different scanners at both 1.5 T and 3 T and with different imaging parameters. It is unlikely that these factors influenced our between group differences given that the normalization process performed by FreeSurfer has good test-retest reliability across different scanners/field strengths (Han et al., 2006; Reuter et al., 2012; Pfefferbaum et al., 2012) and our samples were matched for field strength and imaging parameters. Furthermore, the field strength and average voxel size were included as covariates in the classification analysis. However, including data from multiple scanners surely introduced additional variability into our sample that may have decreased our power to separate the two groups. Yet, it also made the results more generalizable to a real world situation and relevant to future work that might seek to combine data sets across multiple hospitals in order to develop an even more robust and accurate classifier.

Overall, this study gives us promise that future work using larger sample sizes or more homogenous data sets will allow for an even higher accuracy. Furthermore, as we only utilized T1 MRI data and did not incorporate signal variation from T2/FLAIR scans (Coan et al., 2014, Cantor-Rivera et al., 2015) or alterations found in PET (Kerr

et al., 2013) or DTI data (Cantor-Rivera et al., 2015), it is clear that there is still much room for improving classifier performance. Thus, in the future, an even more comprehensive model that identifies subgroups, incorporates other imaging modalities in addition to clinical and neuropsychological data, would likely be superior in regards to both diagnosis and predicting treatment outcomes, thus achieving clinical utility.

5. Conclusions

Beyond MTS as identified through visual inspection by neuroradiologists, there exist no distinct MRI markers used in clinical practice that can be used for diagnosis or predicting treatment response in epilepsy. In this study we developed an automated machine learning approach that was able to identify MTS with up to 81% accuracy in a large sample of epilepsy patients. In addition to hippocampal volume, the algorithm incorporated cortical thickness, surface area and curvature of nearby frontal, temporal and limbic structures to generate classification accuracy. The tendency for an individual to be identified with MTS was associated with earlier age of onset, a longer duration of disease and more frequent seizures. Our findings provide a novel marker in epilepsy by incorporating extra-hippocampal involvement of MTS into a predictive algorithm that may allow clinicians to detect more subtle cases of TLE and MTS. Although still not at the level needed for clinical utility, we hope this line of work will aid in the eventual development of automated neuroimaging-based classification algorithms that will assist in the global neurocognitive evaluation of epilepsy patients, guide treatment decisions, and help predict surgical and overall clinical outcomes.

6. Conflict of interest

The authors report no conflicts of interest.

Acknowledgements

This work was supported by the Ronald Reagan Medical Center, the Department of Radiology at UCLA Center for Health Sciences, the David Geffen School of Medicine at UCLA, and the UCLA Medical Scientist Training Program (T32 GM008044; J.D.R, J.B.C)

References

Alhusaini, S., Doherty, C.P., Palaniyappan, L., Scanlon, C., Maguire, S., Brennan, P., Delanty, N., Fitzsimons, M., Cavalleri, G.L., 2012. Asymmetric cortical surface area and morphology changes in mesial temporal lobe epilepsy with hippocampal sclerosis. Epilepsia 53 (6), 995–1003.

Ambroise, C., McLachlan, G.J., 2002. Selection bias in gene extraction on the basis of microarray gene-expression data. Proc. Natl. Acad. Sci. U. S. A. 99 (10), 6562–6566.

Babb, T.L., Lieb, J.P., Brown, W.J., Pretorius, J., Crandall, P.H., 1984. Distribution of pyramidal cell density and hyperexcitability in the epileptic human hippocampal formation. Epilepsia 25 (6), 721–728.

Berkovic, S.F., Andermann, F., Olivier, A., Ethier, R., Melanson, D., Robitaille, Y., Kuzniecky, R., Peters, T., Feindel, W., 1991. Hippocampal sclerosis in temporal lobe epilepsy demonstrated by magnetic resonance imaging. Ann. Neurol. 29 (2), 175–182.

Bernhardt, B.C., Bernasconi, N., Concha, L., Bernasconi, A., 2010. Cortical thickness analysis in temporal lobe epilepsy: reproducibility and relation to outcome. Neurology 74 (22), 1776–1784.

Bernhardt, B.C., Hong, S.J., Bernasconi, A., Bernasconi, N., 2015. Magnetic resonance imaging pattern learning in temporal lobe epilepsy: classification and prognostics. Ann. Neurol. 77 (3), 436–446.

Cantor-Rivera, D., Khan, A.R., Goubran, M., Mirsattari, S.M., Peters, T.M., 2015. Detection of temporal lobe epilepsy using support vector machines in multi-parametric quantitative MR imaging. Comput. Med. Imaging Graph. 41, 14–28.

Coan, A.C., Kubota, B., Bergo, F.P., Campos, B.M., Cendes, F., 2014. 3T MRI quantification of hippocampal volume and signal in mesial temporal lobe epilepsy improves detection of hippocampal sclerosis. Am. J. Neuroradiol. 35 (1), 77–83.

Colby, J.B., Rudie, J.D., Brown, J.A., Douglas, P.K., Cohen, M.S., Shehzad, Z., 2012. Insights into multimodal imaging classification of ADHD. Front. Syst. Neurosci. 6, p59.

Craddock, R.C., Holtzheimer, P.E., Hu, X.P., Mayberg, H.S., 2009. Disease state prediction from resting state functional connectivity. Magn. Reson. Med. 62 (6), 1619–1628.

De Martino, F., Valente, G., Staeren, N., Ashburner, J., Goebel, R., Formisano, E., 2008. Combining multivariate voxel selection and support vector machines for mapping and classification of fMRI spatial patterns. NeuroImage 43 (1), 44–58.

Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., Killiany, R.J., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage 31 (3), 968–980.

Ecker, C., Rocha-Rego, V., Johnston, P., Mourao-Miranda, J., Marquand, A., Daly, E.M., Brammer, M.J., Murphy, C., Murphy, D.G., MRC AIMS Consortium, 2010. Investigating the predictive value of whole-brain structural MR scans in autism: a pattern classification approach. NeuroImage 49 (1), 44–56.

Engel, J., 1996. Introduction to temporal lobe epilepsy. Epilepsy Res. 26, 141–150.

Farahat, A.K., Ghodsi, A., Kamel, M.S., 2011. An efficient greedy method for unsupervised feature selection. IEEE (Vancouver, BC), 161–170.

Fischl, B., 2004. Automatically parcellating the human cerebral cortex. Cereb. Cortex (New York, N.Y.: 1991) 14 (1), 11–22.

Fischl, B., Dale, A.M., 2000. Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proc. Natl. Acad. Sci. U. S. A. 97 (20), 11050–11055.

Focke, N.K., Yogarajah, M., Symms, M.R., Gruber, O., Paulus, W., Duncan, J.S., 2012. Automated MR image classification in temporal lobe epilepsy. NeuroImage 59 (1), 356–362.

Guyon, I., Weston, J., Barnhill, S., Vapnik, V., 2002. Gene selection for cancer classification using support vector machines. Machine Learn. 46 (1–3), 389–422.

Han, X., Jovicich, J., Salat, D., van der Kouwe, A., Quinn, B., Czanner, S., Busa, E., Pacheco, J., Albert, M., Killiany, R., Maguire, P., Rosas, D., Makris, N., Dale, A., Dickerson, B., Fischl, B., 2006. Reliability of MRI-derived measurements of human cerebral cortical thickness: the effects of field strength, scanner upgrade and manufacturer. NeuroImage 32 (1), 180–194.

Hermann, B., Seidenberg, M., Bell, B., Rutecki, P., Sheth, R.A.J.D., Wendt, G., O'leary, D., Magnotta, V., 2003. Extratemporal quantitative MR volumetrics and neuropsychological status in temporal lobe epilepsy. J. Int. Neuropsychol. Soc. 9 (03), 353–362.

Keller, S.S., Roberts, N., 2008. Voxel-based morphometry of temporal lobe epilepsy: an introduction and review of the literature. Epilepsia 49 (5), 741–757.

Kemmotsu, N., Girard, H.M., Bernhardt, B.C., Bonilha, L., Lin, J.J., Tecoma, E.S., Iragui, V.J., Hagler, D.J., Halgren, E., McDonald, C.R., 2011. MRI analysis in temporal lobe epilepsy: cortical thinning and white matter disruptions are related to side of seizure onset. Epilepsia 52 (12), 2257–2266.

Kerr, W.T., Nguyen, S.T., Cho, A.Y., Lau, E.P., Silverman, D.H., Douglas, P.K., Reddy, N.M., Anderson, A., Bramen, J., Salamon, N., Stern, J.M., Cohen, M.S., 2013. Computer-aided diagnosis and localization of lateralized temporal lobe epilepsy using interictal FDG-PET. Front. Neurol. 4, p31.

Klöppel, S., Stonnington, C.M., Chu, C., Draganski, B., Scahill, R.I., Rohrer, J.D., Fox, N.C., Jack, C.R., Ashburner, J., Frackowiak, R.S., 2008. Automatic classification of MR scans in Alzheimer's disease. Brain 131 (Pt 3), 681–689.

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

Lin, J.J., Salamon, N., Lee, A.D., Dutton, R.A., Geaga, J.A., Hayashi, K.M., Luders, E., Toga, A.W., Engel, J., Thompson, P.M., 2007. Reduced neocortical thickness and complexity mapped in mesial temporal lobe epilepsy with hippocampal sclerosis. Cereb. Cortex (New York, N.Y.: 1991) 17 (9), 2007–2018.

McDonald, C.R., Hagler, D.J., Ahmadi, M.E., Tecoma, E., Iragui, V., Gharapetian, L., Dale, A.M., Halgren, E., 2008. Regional neocortical thinning in mesial temporal lobe epilepsy. Epilepsia 49 (5), 794–803.

Oyegbile, T.O., Dow, C., Jones, J., Bell, B., Rutecki, P., Sheth, R., Seidenberg, M., Hermann, B.P., 2004. The nature and course of neuropsychological morbidity in chronic temporal lobe epilepsy. Neurology 62 (10), 1736–1742.

Pfefferbaum, A., Rohlfing, T., Rosenbloom, M.J., Sullivan, E.V., 2012. Combining atlas-based parcellation of regional brain data acquired across scanners at 1.5 T and 3.0 T field strengths. NeuroImage 60 (2), 940–951.

Téllez-Zenteno, J.F., Hernández Ronquillo, L., Moien-Afshari, F., Wiebe, S., 2010. Surgical outcomes in lesional and non-lesional epilepsy: a systematic review and meta-analysis. Epilepsy Res. 89 (2–3), 310–318.

Voets, N.L., Bernhardt, B.C., Kim, H., Yoon, U., Bernasconi, N., 2011. Increased temporolimbic cortical folding complexity in temporal lobe epilepsy. Neurology 76 (2), 138–144.