NeuroImage: Clinical 9 (2015) 20–31

Contents lists available at ScienceDirect

NeuroImage: Clinical

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

Predicting the laterality of temporal lobe epilepsy from PET, MRI, and DTI: A multimodal study

Predicting temporal lobe epilepsy laterality

Dorian Pustina [a], Brian Avants [b], Michael Sperling [c], Richard Gorniak [d], Xiaosong He [c], Gaelle Doucet [c], Paul Barnett [c], Scott Mintzer [c], Ashwini Sharan [e], Joseph Tracy [c,d,*]

[a] Department of Neurology, University of Pennsylvania, Philadelphia, USA
[b] Department of Radiology, University of Pennsylvania, Philadelphia, USA
[c] Department of Neurology, Thomas Jefferson University/Sidney Kimmel Medical College, Philadelphia, PA 19107, USA
[d] Department of Radiology, Thomas Jefferson University/Sidney Kimmel Medical College, Philadelphia, PA 19107, USA
[e] Department of Neurosurgery, Thomas Jefferson University, Philadelphia, USA

ARTICLE INFO

Article history:
Received 14 May 2015
Received in revised form 11 July 2015
Accepted 19 July 2015
Available online 31 July 2015

Keywords:
Asymmetry
Classification
Metabolism
Resection
Machine learning

ABSTRACT

Pre-surgical evaluation of patients with temporal lobe epilepsy (TLE) relies on information obtained from multiple neuroimaging modalities. The relationship between modalities and their combined power in predicting the seizure focus is currently unknown. We investigated asymmetries from three different modalities, PET (glucose metabolism), MRI (cortical thickness), and diffusion tensor imaging (DTI; white matter anisotropy) in 28 left and 30 right TLE patients (LTLE and RTLE). Stepwise logistic regression models were built from each modality separately and from all three combined, while bootstrapped methods and split-sample validation verified the robustness of predictions. Among all multimodal asymmetries, three PET asymmetries formed the best predictive model (100% success in full sample, >95% success in split-sample validation). The combinations of PET with other modalities did not perform better than PET alone. Probabilistic classifications were obtained for new clinical cases, which showed correct lateralization for 7/7 new TLE patients (100%) and for 4/5 operated patients with discordant or non-informative PET reports (80%). Metabolism showed closer relationship with white matter in LTLE and closer relationship with gray matter in RTLE. Our data suggest that metabolism is a powerful modality that can predict seizure laterality with high accuracy, and offers high value for automated predictive models. The side of epileptogenic focus can affect the relationship of metabolism with brain structure. The data and tools necessary to obtain classifications for new TLE patients are made publicly available.

1. Introduction

Epilepsy is a neurologic disorder that affects ~70 million people worldwide (Ngugi et al., 2010). Typically, seizures originate in a focal brain area, called the seizure onset zone (SOZ), and epilepsies are eventually classified based on the location of SOZ in the brain. The most common type of epilepsy is temporal lobe epilepsy, accounting for about 65% of focal epilepsy cases (TLE; Engel, 2001; Tellez-Zenteno and Hernandez-Ronquillo, 2012). About one third of TLE patients do not respond to medication, and become candidates for surgical resection of the epileptogenic tissue.

The success of surgery depends on the correct identification of the SOZ in each patient. Besides EEG and neuropsychologic evaluation, a repertoire of neuroimaging tools are frequently used to improve SOZ detection, among which MRI, PET, and diffusion imaging (Bonilha and Keller, 2015; Sperling et al., 1992; Winston, 2015). The various neuroimaging modalities are reviewed by expert radiologists who visually inspect the images in search of pathology. When found, asymmetric pathologies are a good indicator of the presence of the SOZ, providing a good prognosis for improved seizure control once the pathologic area is resected (Blum et al., 1998; Englot and Chang, 2014; Gok et al., 2013; LoPinto-Khoury et al., 2012; Tonini et al., 2004). The most common pathologies observed in TLE are: an asymmetric reduction of glucose metabolism (known as hypometabolism) in one of the temporal lobes, an asymmetric atrophy of mesial temporal lobe structures (known as mesial temporal, or hippocampal, sclerosis), or other local morphologic anomalies such as dysplasias and heterotopias. Besides pathologies related to gray matter (GM), white matter (WM) anomalies

* Corresponding author at: Departments of Neurology and Radiology, Thomas Jefferson University/Sidney Kimmel Medical College, 900 Walnut Street, Health Sciences Building, Suite 447, Philadelphia, PA 19107, USA. Tel.: +1 (215)955 4661; fax: +1 (215)955 3745.
E-mail address: joseph.tracy@jefferson.edu (J. Tracy).

http://dx.doi.org/10.1016/j.nicl.2015.07.010

have also been reported in TLE, some of which are asymmetric and can help localize the SOZ (Ahmadi et al., 2009; Concha et al., 2012; Thivard et al., 2011).

The ability to identify pathologies from neuroimaging data requires many years of training and additional dedicated time in routine clinical practice. Even with the right training and expertise, the accuracy of human experts may depend on other subjective factors, such as the amount of clinical knowledge the expert has about the case (Thivard et al., 2011). Recent developments in neuroimage analysis have shown that automated quantification and prediction can be a powerful and promising tool in medical practice (for a review, see Gabrieli et al., 2015). When applied to epilepsy, early studies showed that automated neuroimage analysis and patient classification can match or exceed the accuracy of visual inspection of human experts (Ferrie et al., 1997; Lee et al., 2000; Matheja et al., 1998). More recently, several research papers have built predictive models that can distinguish among subtypes of epilepsy patients based on PET data (Didelot et al., 2010; Lee et al., 2000; Muzik et al., 1998; Soma et al., 2012), MRI data (Duchesne et al., 2006; Moser et al., 2000), or a combination of PET and MRI (Lopez-Acevedo et al., 2012). The power of DTI has also been investigated for automated classification either as a single modality (Ahmadi et al., 2009; Concha et al., 2012) or in combination with PET and MRI (Thivard et al., 2011). This knowledge, however, has not been readily translated into tools of clinical utility. Moreover, the availability of many neuroimaging protocols raises the question to whether the combination of multiple imaging modalities into a single predictive model can improve accuracy rates above and beyond what is obtained from each modality separately.

The aim of this study was to provide a comprehensive investigation of data obtained from three imaging modalities with the intent of creating an automated classification algorithm with clinical applicability, capable of distinguishing left from right temporal lobe epilepsy (LTLE and RTLE, respectively). We aimed at creating a predictive tool that does not simply classify patients, but also provides a probabilistic index of the classification. In addition, we investigated whether the relationship between GM, WM, and metabolism, is the same in LTLE and RTLE patients. We utilized asymmetry measures from all modalities for three main reasons: (i) human experts frequently look at degrees of asymmetry when inspecting neuroimaging data, (ii) asymmetries use each subject as a control for itself, thus decreasing confounds coming from differences in scanner parameters, time of scan, metabolic factors, etc., and (iii) asymmetries have been shown to be more sensitive to the pathology than raw voxel values (Didelot et al., 2010; Soma et al., 2012; Van Bogaert et al., 2000).

2. Material and methods

2.1. Subjects

A total of 58 patients with refractory, unilateral and focal temporal lobe epilepsy were recruited from the Thomas Jefferson University Comprehensive Epilepsy Center (28 left TLE; 30 right TLE). All patients had solely unilateral temporal pathology, and all were evaluated for surgical resection of the epileptogenic area as treatment for their intractable temporal lobe epilepsy. Details of the Thomas Jefferson Comprehensive Epilepsy Center algorithm for surgical decision making are described in Sperling et al. (1992). A combination of EEG (at least 96 h), video recording, MRI, PET, and neuropsychological testing, was used to localize the seizure focus. Expert board-certified neuroradiologists, epileptologists, and neuropsychologists, by consensus decision, classified patient neuropathology and seizure type. When the data from the various tests was not sufficiently informative, intracranial electrodes were implanted to determine the seizure focus through electrocorticography (12 patients, 5 LTLE).

Participants were excluded if they were missing data from one of the modalities (MRI, DTI, or PET). Other exclusion criteria consisted of medical illness with central nervous system impact other than epilepsy; prior or current alcohol or illicit drug abuse; extratemporal epilepsy; present or past neoplasia; contraindications to MRI; psychiatric diagnosis other than an Axis-I Depressive Disorder; or hospitalization for any Axis I disorder listed in the Diagnostic and Statistical Manual of Mental Disorders, IV. Depressive Disorders were allowed in the patient sample, given the high co-morbidity of depression and epilepsy (Tracy et al., 2007). Patients with mental retardation (Full-Scale IQ < 70) who were likely to be unable to cooperate with the MRI examination were also excluded. Patients with bilateral mesial temporal sclerosis were also excluded. The study was approved by the Thomas Jefferson University Institutional Review Board for Research with Human Subjects, and consent was obtained from all patients. Table 1 displays the mean values and frequencies for the relevant demographic and clinical characteristics of our sample.

2.2. Neuroimaging protocols

2.2.1. MRI

Images were acquired in a Philips Achieva 3T scanner (Amsterdam, The Netherlands) using an 8-channel SENSE head coil. The MP-RAGE and the DTI sequence were acquired during the same scanning session. No single group was scanned on a schedule or time different than the other groups, thereby avoiding bias related to temporally dependent scanner calibration. The MP-RAGE volume was collected in sagittal orientation with in-plane resolution of 256×256 and 1 mm slice thickness (isotropic voxels of 1 mm^3; TR = 650 ms, TE = 3.2 ms, FOV 256 mm, flip angle 8°, SENSE factor = 1).

The diffusion data were acquired using a single-shot spin-echo EPI pulse sequence (TE = 90 ms, TR = 8609 ms, SENSE factor = 2.5, 5 min acquisition) with 32 diffusion weighted (b-factor = 850 s/mm^2, anterior–posterior fold-over direction) and three non-diffusion volumes (b-factor = 0 s/mm^2; averaged in scanner to a single B0 volume). Each volume contained 66 slices (thickness = 2 mm, gap = 0 mm) acquired in the axial plane, with a reconstructed matrix size of 128×128 and FOV 230 mm, resulting in voxel size $1.8 \times 1.8 \times 2$ mm. Fat suppression was achieved using a standard SPIR (spectral pre-saturation with inversion recovery) technique. The sequence was repeated twice for each subject and the datasets were averaged to increase signal to noise ratio.

2.2.2. FDG-PET

All PET scans were performed during interictal periods using a standard protocol.

Pre-injection blood glucose level was below 150 mg/dl for all patients (range: 61–128 mg/dl). An intravenous catheter was inserted under local anesthesia and a dose around 5.9 mCi (STD 1.42) of radioactive 100 mg/l fludeoxyglucose (FDG) was injected. The scan was initiated about 42 min (STD 14.8) after the injection. The eyes were open, the ears were non-occluded, and ambient noise and light was kept to a minimum. Forty-seven patients (81% of patients; 24 LTLE, 23 RTLE) were scanned on a Siemens Biograph 1080 camera (Siemens Medical Solutions, Erlangen, Germany), with data consisting of 109 axial slices, 3 mm thick, and 1×1 mm in resolution (scan dates 2008–2013). The remaining 11 subjects were scanned on two different cameras (scan dates 2006–2008), a Siemens Biograph CPS 1080 with voxel size 1×1 mm (109 slices, 2 mm thick), and a Philips Allegro PET Body (Philips Medical Systems, Best, The Netherlands) with voxel size 2×2 mm (90 slices, 2 mm thick). Note, we obtained asymmetry indices that use the same subject as reference, therefore reducing potential scanner specific bias. This procedure also avoids confounds related to demographic factors, such as, age, medication history, epilepsy history, etc. (Leiderman et al., 1991; Theodore, 1989). Nevertheless, a cross validation of results between different scanners was performed (Results section).

Table 1
Values indicate average ± standard deviation. Abbreviations: MTS = mesial temporal sclerosis, MTL = mesial temporal lobe, CPS = complex partial seizures, GTCS = generalized tonic-clonic seizures, SPS = simple partial seizures, w/2° GTCS = with secondary generalization, rare GTCS = ten or less GTCS episodes in lifetime.

	LTLE	RTLE
N	28	30
Age	42.4 ± 12.1	39.6 ± 14.9
Gender F/M	15/13	13/17
Intracranial volume (l)	1.39 ± 0.23	1.43 ± 0.25
Edinburgh handedness	84 ± 41.4	58 ± 65.1
Full Scale IQ	95.8 ± 14.4	95.1 ± 13.1
Age at epilepsy onset	22.3 ± 15.1	21.0 ± 12.0
Duration of epilepsy (yrs)	20.1 ± 15.3	18.6 ± 15.1
MTS/non-MTS (no. of patients)	12/16	10/20
MRI-negative	9 (3 implanted)	14 (4 implanted)
Non-MTS MRI pathology	1 — L cavernous hemangioma in fusiform g. (10 mm)	1 — enlarged R temporal horn
	1 — L ant. TL cavernoma (15 mm)	1 — subtle R hippo signal increase
	1 — signal increase in L hippocampus, no atrophy (implanted)	1 — R posteromedial TL dysplasia
	1 — L TL dysplasia	1 — R TL atrophy (implanted)
	1 — L TL dysplasia and low grade glioma (10 mm)	1 — R > L hippo atrophy, normal MRI signal
	1 — hyperintensity in L MTL, mild amygdala atrophy	1 — R TL white matter FLAIR abnormality (implanted)
	1 — meningoencephalocele in L middle cranial fossa	
Seizure types (no. of patients)	12 — CPS only	13 — CPS only
	5 — CPS + GTCS	3 — CPS + GTCS
	3 — CPS + rare GTCS	2 — CPS + rare GTCS
	2 — SPS only	1 — SPS only
	2 — CPS w/2° GTCS	5 — CPS w/2° GTCS
	3 — CPS + SPS	5 — CPS + SPS
	1 — CPS w/2° GTCS + SPS	1 — CPS w/2° GTCS + SPS
Medication (no. of patients)	6 — Carbamazepine	3 — Carbamazepine
	5 — Depakote	5 — Depakote
	5 — Lacosamide	4 — Lacosamide
	5 — Lamotrigine	5 — Lamotrigine
	12 — Levetiracet	8 — Levetiracet
	1 — Oxcarbazepine	5 — Oxcarbazepine
	1 — Phenobarbital	1 — Phenobarbital
	2 — Phenytoin	2 — Phenytoin
	6 — Topiramate	7 — Topiramate
	0 — Zonisamide	1 — Zonisamide
	2 — Other	2 — Other

2.3. Neuroimaging analyses

2.3.1. MRI — cortical thickness asymmetries

The rationale for considering cortical thickness as a potential predictor of SOZ laterality followed the reports of several groups who found more widespread atrophy in LTLE than in RTLE (Bonilha et al., 2007; Keller and Roberts, 2008; Kemmotsu et al., 2011).

Two popular software were used to estimate cortical thickness in order to cross-validate potential findings and to investigate differences arising from estimation methods: Freesurfer v.5.3.0 (http://freesurfer.net; Fischl and Dale, 2000) and Advanced Normalization Tools v.2.1.0 (ANTs, http://stnava.github.io/ANTs; Avants et al., 2011). In Freesurfer, a fully automated thickness estimation was obtained from the MP-RAGE volume using default parameters and the 'recon-all' command with the Desikan atlas as a parcellation reference (Desikan et al., 2006). Surface maps were visually inspected for major failures, and none were found. Average thickness was then obtained from the lateral temporal lobe (superior, middle, and inferior temporal gyri; parcels 1030/1015/1009 for left and 2030/2015/2009 for right) and the mesial temporal lobe (parahippocampal gyrus and parahippocampal cortex; parcels 1030/1015/1009 for left and 2030/2015/2009 for right). The parcels were combined using a weighted mean depending on area size; parcels with larger surface area contributed more to the average. The hippocampal volume constituted a third measure of interest. Asymmetries were computed with the formula (right − left) / (right + left). The Freesurfer lateral thickness, mesial thickness, and hippocampal volume were labeled "THICK-lat-Frees", "THICK-mes-Frees", and "HIPPO-vol".

In ANTs, thickness maps were obtained using an established pipeline that included bias correction, multivariate segmentation, and thickness estimation with the DIREcT algorithm (all incorporated in the script antsCorticalThickness.sh provided with the software; Tustison et al., 2014). Note that ANTs uses a voxel based method of thickness estimation, while Freesurfer uses a surface based method (Avants et al., 2011; Klein et al., 2010). After obtaining the thickness map with default parameters, which weighted segmentations priors at $\alpha = 0.25$, regional parcellation was performed using a multi-atlas parcellation algorithm. This procedure calculates the parcellation based on the similarity of multiple atlases (NKI-TRT-20 dataset, http://mindboggle.info) with the patient's brain ('antsMalfLabeling.sh'; Wang et al., 2012). Thickness values for lateral and medial temporal lobe were obtained in ANTs from the same regions used in Freesurfer; only voxels that were classified as gray matter during segmentation were included in the computation. While hippocampal volume was not available in ANTs, we obtained an additional measure of entire temporal lobe asymmetry by combining mesial and lateral temporal regions. This measure matched a similar PET measure obtained from the entire temporal lobe (see Section 2.3.3 PET — metabolism asymmetries). The final ANTs variables were labeled "THICK-lat-Ants", "THICK-mes-Ants", and "THICK-entire-Ants".

2.3.2. DTI — tract asymmetries

Prior to processing, diffusion data were visually inspected to identify and potentially remove volumes with excessive motion; no such volumes were found. Subject motion, eddy currents, and EPI distortions were corrected in a single step using cubic spline interpolation in ExploreDTI (Leemans et al., 2009). Tensor estimation was performed with the non-linear RESTORE algorithm, a procedure that is robust to outlier data points (Chang et al., 2005; Chang et al., 2012). Maps of fractional anisotropy (FA) were exported for later computation of tract

asymmetries. Data preprocessed in ExploreDTI were imported in FSL for tractography reconstruction (Diffusion Toolbox, www.fmrib.ox.ac.uk/fsl; Behrens et al., 2007). A Monte Carlo sampling procedure initially estimated the diffusion properties in each voxel, accounting for two tensors per voxel ('bedpostx' command). Afterward, a probabilistic algorithm was used ('probtrax2' command) to initiate 5000 fiber samples per seed voxel with the default angular threshold of 78°. The first author (DP) manually drew seed and waypoint masks using anatomical landmarks visible in the colorized FA map. After obtaining the tractography map, voxels with less than 10% of the maximum visitation count were set to zero to remove spurious streamlines, and a visual inspection in 3D was performed in search of potential failures (see DTI tractography examples in Supplementary Information; see also Pustina et al., 2014b for another example of this pipeline).

In total, eight tracts were reconstructed for each subject, four in each hemisphere: (i) the fornix, (ii) the parahippocampal fasciculus, (iii) the uncinate fasciculus, and (iv) the temporo-occipital connections consisting mostly of the inferior longitudinal fasciculus. For each tract, the FA value was computed as a weighted average of the probability mask, i.e. voxels with high visitation counts influenced the average more than voxels with low visitation counts. Asymmetries were calculated with the formula (right − left) / (right + left). The final DTI asymmetries were labeled "DTI-FA-fx", "DTI-FA-phip", "DTI-FA-unc", and "DTI-FA-ilf".

2.3.3. PET — metabolism asymmetries

The PET volume was coregistered to the respective MP-RAGE volume of each subject using SPM8 (University College London, http://www.fil.ion.ucl.ac.uk/spm/software/spm8/). Four regional masks were utilized to extract PET values for the lateral, mesial, and entire temporal lobe, as well as the hippocampus. The masks were obtained from the Freesurfer parcellation map and contained, respectively, (i) the gray matter of the lateral temporal lobe (parcels 1030/1015/1009 for left and 2030/2015/2009 for right), (ii) the gray matter of mesial temporal lobe including amygdala and hippocampus (parcels 17/18/1006/1016 for left and 53/54/2006/2016 for right), (iii) only the hippocampus (parcel 17 for left and 53 for right hemisphere), and (iv) the mesial and lateral masks combined together. Note, the four regional masks were partially overlapping in order to determine which combination of lateral and mesial areas yielded the best discriminatory power. During the preliminary phase of the study we noticed that the cortical mantel zigzagged around the PET blobs, and the PET values obtained with such masks were unreliable. We resolved this issue by processing the masks to reflect the smoothness of PET data (see example in Supplementary Information: PET masking). To understand this procedure, consider that the metabolic signal is emanated from glucose uptake occurring mostly in gray matter, which is represented by the binary masks obtained from Freesurfer. This signal, however, is distorted due to limitations of PET imaging (i.e., the finite size of detector crystals, detector physics, travel distance before annihilation, Poisson count statistics, and reconstruction methods). The final PET image can be theoretically approximated by assuming that the true metabolic signal is smoothed with a Gaussian smoothing kernel (the point spread function of the scanner; Greve et al., 2014). We can apply a similar smoothing to the cortical masks. Once smoothed, values will remain closer to 1 in areas of densely packed GM, and will fade closer to 0 in areas of thin GM. The metabolic signal undergoes a similar distortion such that PET values near densely packed GM are closer to the true metabolic signal, while PET values of thinner GM are farer from the true metabolic signal. Thus, we applied 8 mm FWHM smoothing to the binary masks, and used the resulting continuous masks as weights to compute weighted average of PET for each region. Note, the PET volume itself was not smoothed, only the mask was smoothed. Voxels below 0.35 in the continuous mask were set to 0 to minimize the influence of non-GM signal. The final asymmetry values were computed with the formula: (right − left) / (right + left). The four PET asymmetries were labeled "PET-hippo", "PET-entire", "PET-lateral", and "PET-mesial".

2.4. Extended dataset: variance, skewness and kurtosis

In epilepsy, the seizure focus may manifest as a small region with low metabolism, abnormal GM thickness or volume, or abnormal WM

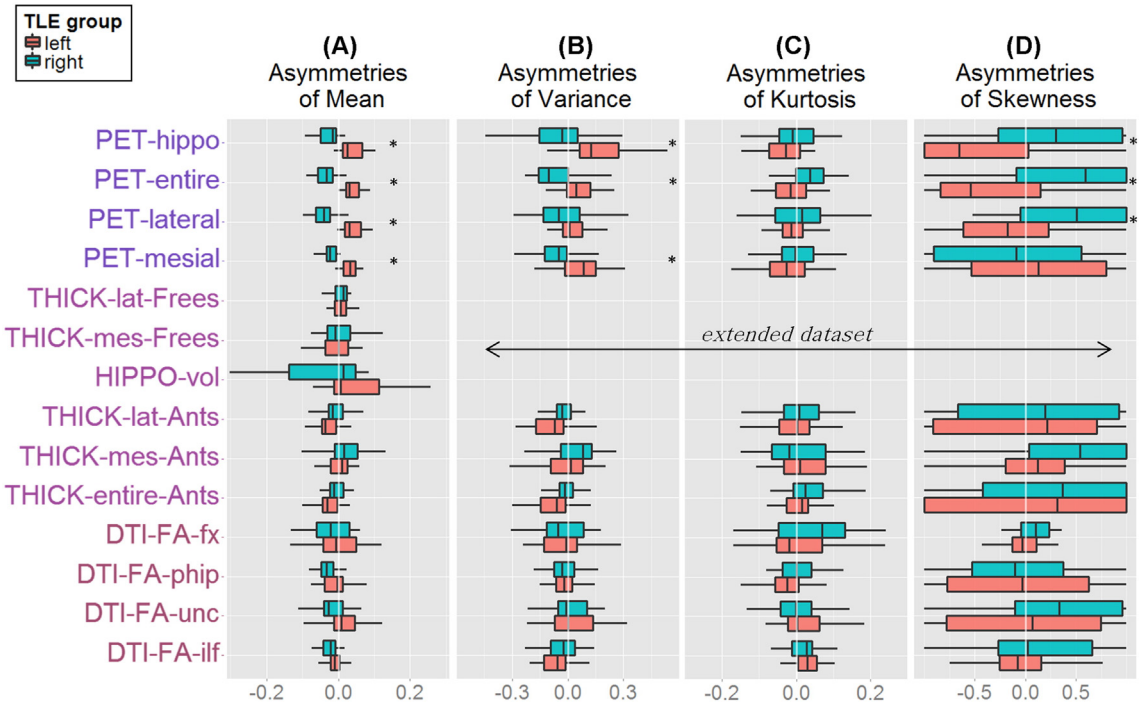

Fig. 1. Multimodal asymmetries of the mean (A), variance (B), kurtosis (C), and skewness (D) in LTLE (red boxes) and RTLE (green boxes). Boxplots extend between 25th and 75th percentile of the data, the central line indicates the median value. Whiskers extend to the most extreme non-outlier data point (outliers values = 1.5 × interquartile range value). The main dataset consists in data from A, the extended dataset consists in data from A–B–C–D. Asterisks mark significant group differences (Wilcoxon tests) after multiple comparison correction.

structure. This pattern may not be detected when averaging the signal from large regional masks. Instead, local anomalies may manifest as increased variance in the data, or skewed and abnormally distributed data. We investigated the value of adding variables that describe the shape of the data distribution by creating an extended dataset, where each asymmetry involving a mean was complemented with additional asymmetries of variance, skewness, and kurtosis (Fig. 1). For PET and DTI, which had probability masks, a weighted variance algorithm was applied (variance in regions of high probability counted more than variance in regions of low probability). All the other extended dataset asymmetries were computed from binarized masks. The asymmetries of skewness and kurtosis were computed with a modified formula to account for the presence of negative values: (right − left) / abs(right) + abs(left).

2.5. Statistical analyses

Analyses were performed using SPSS (IBM SPSS Statistics for Windows v.20, Armonk, NY), and R (R Foundation for Statistical Computing, http://www.r-project.org). Demographic profiles (i.e., age, IQ, scan intervals, etc.) were compared between left and right TLE patients with t-test or chi-square, as appropriate, and a correction for variance inhomogeneity was applied when necessary.

Shapiro-Wilk tests on neuroimaging variables revealed non-normal distribution on some asymmetries. Consequently, asymmetry differences between groups were compared with the non-parametric Wilcoxon test. The maxT procedure with 100 million permutations was used to correct for multiple comparisons ('multtest' package in R; Pollard et al., 2005). The maxT procedure extracts the dependency structure of statistical tests by running permutations on raw data rather than considering tests as independent and correcting only p-values. In general, maxT is considered more conservative than the false discovery rate and less conservative than Bonferroni (Ge et al., 2003; Westfall and Young, 1993). Corrected p-values were thresholded at $\alpha = 0.05$.

Logistic regression was used to discriminate the left and right TLE groups and predict group membership. In contrast with linear discriminant analysis, logistic regression does not assume normal data distribution or equal variance between groups. A stepwise procedure was used to select relevant variables that can distinguish LTLE from RTLE ("forward conditional" option in SPSS with entry cutoff of $p < 0.05$). After building the regression models, their reliability was tested with bootstrapped and split-sample procedures. For bootstraps run on the full sample, the group size was kept identical to the original dataset (28 and 30, respectively), while 10,000 random samples were drawn with replacement. For bootstraps run in split samples, the data were randomly split into training and testing groups without replacement. The size of the training group varied progressively from 97% to 52%, reducing the size in steps of 3% (approximately removing one patient per group in 16 steps). The process was repeated 10,000 times at each step to provide a reliable estimate of the success rates of the model with unseen data. All bootstraps were performed with a penalized regression algorithm, which avoids infinite beta values when complete separation occurs (Firth, 1993). The 'brglm' package in R was used for this purpose (Kosmidis, 2013). The output of logistic regression models is in the range 0–1, with 0 representing LTLE and 1 representing RTLE. The threshold for classifying patients was set to 0.5. Two variants of predictive models were attempted: using each modality separately or using all three modalities. When combined modalities were utilized, the candidate predictors were either all the 14 available asymmetries of the main dataset, or those that were significantly different between the two groups after multiple comparison correction (to be referred to as the unthresholded and thresholded models, respectively). A similar procedure was applied with the extended dataset which had 47 available asymmetries.

Differences in average inter-modality correlations between groups were investigated with independent t-tests. For this purpose, all correlation coefficients between a modality pair (i.e., PET-DTI) obtained from one group were compared with those of the other group (see Pustina et al., 2015 for a previous application). In total, three t-tests were run, one for each modality pair, with results thresholded at $\alpha = 0.05/3 = 0.0167$ to correct for multiple comparisons. A qualitative inspection of the links between asymmetries was also performed using dendrogram trees from bootstrapped hierarchical clustering (Suzuki and Shimodaira, 2006).

3. Results

All 58 patients included in this study were PET-positive (a pathology was detected in PET by expert visual inspection), while 35 of them were also MRI-positive (a pathology was detected in MRI by expert visual inspection). The lack of PET-negative patients derived from the need for data in all three modalities and the other exclusion criteria. The most common pathology found in MRI was MTS (22 patients, 40% of all patients, 63% of MRI-positive patients). The typical pathology found in PET was hypometabolism in one of the temporal lobes, none of the patients had hypermetabolism in the temporal lobe. Relevant demographic and clinical information is shown in Table 1. The female/male ratio was 15/13 in the LTLE group and 13/17 in the RTLE group. The two TLE groups did not differ in age ($t[55] = .767$, $p = .446$), gender (chi-square = .608, $p = .436$), handedness ($t[50] = 1.98$, $p = .070$), age of epilepsy onset ($t[56] = .903$, $p = .717$), epilepsy duration ($t[56] = .903$, $p = .717$), or intracranial volume ($t[56] = -.554$, $p = .582$).

3.1. Asymmetry differences among TLE groups

Fig. 1 displays the boxplots of each asymmetry variable in LTLE and RTLE. After maxT correction for multiple comparisons across the 14 main dataset asymmetries (Fig. 1A), none of the DTI or MRI asymmetries survived the correction, while all four PET asymmetries were significantly different between LTLE and RTLE. The extended dataset (including variance, skewness, and kurtosis) showed ten asymmetries to be different between groups after maxT correction, all of which were derived from PET: PET-mesial ($W = -6.32$, p-adj < 0.001), PET-entire ($W = -6.30$, p-adj <0.001), PET-lateral ($W = -6.18$, p-adj < 0.001), PET-hippo ($W = -5.98$, p-adj < 0.001), PET-entire-variance ($W = -4.01$, p-adj $= 0.001$), PET-hippo-variance ($W = -3.97$, p-adj $= 0.002$), PET-entire-skewness ($W = 3.90$, p-adj $= 0.002$), PET-hippo-skewness ($W = 3.69$, p-adj $= 0.006$), PET-lateral-skewness ($W = 3.44$, p-adj $= 0.016$), and PET-mesial-variance ($W = -3.28$, p-adj $= 0.028$).

3.2. Logistic regression within each modality

The stepwise regression of PET asymmetries selected PET-mesial and PET-lateral as significant predictors of group membership (accuracy = 96.6%, chi-square = 70.95, $p < .001$, Nagelkerke R2 = .941, Hosmer & Lemenshow p-value = 1).

The stepwise regression of MRI asymmetries selected only THICK-entire-Ants as a significant predictor of group membership (accuracy = 67.2%, chi-square = 6.67, $p = .01$, Nagelkerke R2 = .145, Hosmer & Lemenshow p-value = .698). In an attempt to match the variables obtained from Freesurfer and ANTs, we also obtained the same THICK-entire asymmetry from Freesurfer, but a t-test showed no difference between TLE groups ($p > .1$) and the Freesurfer variable did not enter the predictive equation.

The stepwise regression of DTI symmetries selected only DTI-FA-ilf as significant predictor of TLE group membership (accuracy = 70.7%, chi-square = 6.84, $p = .009$, Nagelkerke R2 = .148, Hosmer & Lemenshow p-value = .642).

3.3. Logistic regression using all multimodal asymmetries

A stepwise regression with all 14 multimodal asymmetries as candidate predictors selected two PET (PET-mesial and PET-lateral) and one MRI (THICK-lateral-Ants) asymmetries as significant predictors (accuracy = 100%, chi-square = 80.34, $p < .001$, Nagelkerke R2 = 1, Hosmer & Lemenshow p-value = 1).

As mentioned, only PET asymmetries survived multiple comparison adjustment in group comparisons. Therefore, there was no chance for other modalities to enter the model created by thresholded data, and the same model as with PET alone emerged (accuracy = 96.6%, chi-square = 70.95, $p < .001$, Nagelkerke R2 = .941, Hosmer & Lemenshow p-value = 1).

3.4. Logistic regression on the extended dataset

Besides the existing asymmetries of mean, the extended dataset included the asymmetries of variance, skewness and kurtosis, across all three modalities. A stepwise regression with all 47 asymmetries selected four significant predictors: PET-lateral, PET-mesial, DTI-FA-phip, PET-entire-variance (accuracy = 100%, chi-square = 80.35, $p < .001$, Nagelkerke R2 = 1, Hosmer & Lemenshow p-value = 1).

As mentioned, ten asymmetries survived multiple comparison adjustment from the extended dataset, all of which were PET related. A stepwise regression with these asymmetries selected three variables (PET-mesial, PET-hippo-variance, and PET-entire-variance) as the best predictors of the TLE group (accuracy = 100%, chi-square = 80.34, $p < .001$, Nagelkerke R2 = 1, Hosmer & Lemenshow p-value = 1).

3.5. Selection of the most appropriate prediction model

The above analyses showed that three models were the most successful in predicting group membership, all reaching 100% accuracy. The first was obtained from unthresholded data in the main dataset and included two PET and one MRI variable (Model 1), the second was obtained from the thresholded extended dataset and included three PET variables (Model 2), the third was obtained from the unthresholded extended dataset and included three PET and one DTI variable (Model 3). In selecting the most appropriate model we considered three factors: (i) the amount of erroneous classifications from a full-sample bootstrapped procedure, (ii) the accuracy rates obtained from the split-sample validation procedure, and (iii) whether a bias existed in predicting one group better than the other. Success rates from the full-sample bootstrapping procedure were slightly lower for the first model (95% CI [96% 100%]) compared to the other two ([98% 100%] and [100% 100%], respectively). We therefore excluded Model 1 and considered further Models 2 and 3. Their bootstrapped validations are plotted in Fig. 2. With regard to the split-sample validation, both models did surprisingly well, with above 95% accuracy at all steps. However, the bootstrap on the full sample showed a largely unbalanced prediction with the third model, in which LTLE patients carried more certain posterior probability values than RTLE (lower right panel in Fig. 2). On the contrary, posterior probabilities of the second model were similar between LTLE and RTLE patients. Consequently, the second model was deemed as the winning model as it satisfied all three criteria.

3.6. Cross-validation of PET scanners

Given that the winning predictive model was derived from three PET asymmetries, we tested whether data coming from different scanners can be used to predict group membership. For this, we ran a bootstrapped cross-validation with the patients scanned with the main Siemens camera as training group (N = 47, resampling with replacement) and the remaining patients as test group (N = 11). The average success rate was 99.4% with 95% CI = [91% 100%], indicating a

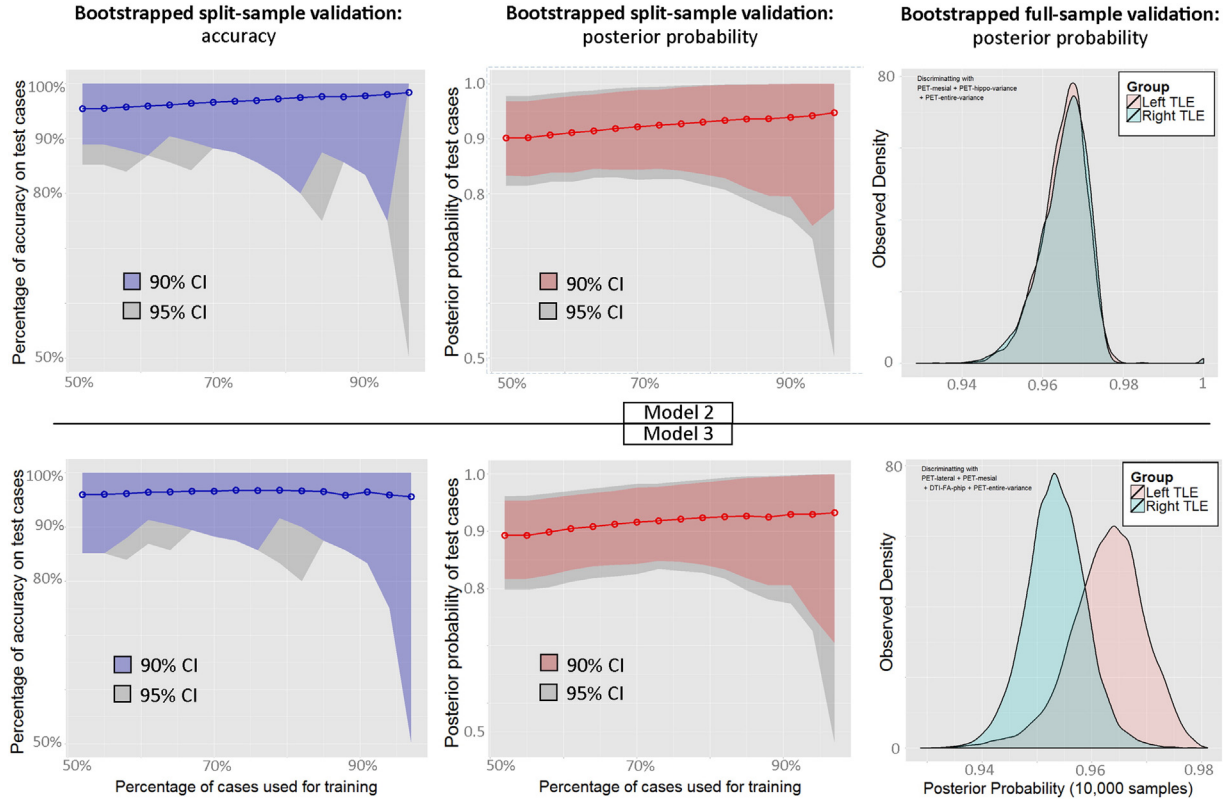

Fig. 2. Bootstrapped validation of model 2 (upper panels) and model 3 (lower panels). Left panels: split-sample validation accuracy. Middle panels: split-sample posterior probability. Right panels: distribution of posterior probability from 10,000 full-sample bootstraps. Note, the posterior probability of LTLE was flipped for comparison with RTLE (i.e., 0.03 became 0.97).

Table 2
The column labeled "PET prediction asymmetries" shows values used as input to the prediction, and correspond to asymmetries of PET-mesial, PET-hippo-var, and PET-entire-var, respectively. Surgeries for "New TLE" were performed during the past 12 months. Abbreviations: L = left, R = right, TL = temporal lobe, MTS = mesial temporal sclerosis, ATL = anterior temporal lobectomy, CPS = complex partial seizures, GTCS = generalized tonic-clonic seizures, SPS = simple partial seizures, w/2° GTCS = with secondary generalization, rare GTCS = ten or less GTCS episodes in lifetime.

ID	Seizure type	Ictal EEG	MRI inspection	PET inspection	Clinical decision	PET prediction asymmetries	Probabilistic classification	95% CI posterior probability	Surgery + outcome
New TLE									
1-LO	CPS, CPS w/2° GTCS	L TL electrodes	L MTS	L anter. TL hypo.	LTLE	0.01553 −0.00439 −0.02010	LTLE 99.72%	[0.027–0.339]	ATL + Class I
2-FU	CPS, rareGTCS	L TL and frontal electrodes	Unremarkable	Bil. TL hypo, biased to left	LTLE	0.04544 0.15939 −0.02126	LTLE 100%	[0–0.01]	n.a.
3-RA	CPS	L frontal electrodes	Unremarkable	L later. TL hypo.	LTLE	0.00434 −0.08219 0.07710	LTLE 99.64%	[0.04–0.36]	Implants + ATL + Class I
4-NE	CPS, GTCS	L frontal electrodes	L MTS	L TL hypo.	LTLE	0.03004 0.33749 0.08012	LTLE 100%	[0–0.05]	Laser ablation + Class I
5-GM	CPS, rare GTCS	R Temporal electrodes	R MTS	R TL hypo.	RTLE	−0.05851 −0.20535 −0.16405	RTLE 100%	[0.99–1]	Laser ablation + Class I
6-OL	SPS, CPS, CPS w/2° GTCS	R TL electrodes	R MTS	R TL hypo.	RTLE	−0.03744 0.05089 0.12806	RTLE 98.47%	[0.63–0.99]	ATL + Class I
7-NL	SPS, CPS	R TL electrodes	General atrophy, larger R amygdala	Mild R mesial TL hypo.	RTLE	−0.02331 0.20569 0.05303	RTLE 96.33%	[0.44–0.99]	ATL + Class I
Ambiguous cases									
1-RD	CPS, rare GTCS	L TL electrodes	Hyperintensity in R frontal, R periventricular, L caudate, and mild R hippocampal atrophy	Bilateral hypo, biased to left	LTLE	0.00729 −0.04469 0.03298	LTLE 99.49%	[0.075–0.389]	Implants + ATL + Class I
2-VO	SPS, CPS	L TL electrodes	L MTS	Bilateral hypo.	LTLE	0.02933 0.05192 −0.04924	LTLE 99.97%	[0.001–0.157]	Laser ablation + Class I
3-VD	CPS	L TL electrodes	L MTS	Bilateral hypo.	LTLE	0.04426 0.17092 0.08998	LTLE 100%	[0–0.004]	ATL + Class I
4-CS	CPS, CPS w/2° GTCS	L TL electrodes	Unremarkable	Unremarkable	LTLE	0.02901 0.12306 0.15600	LTLE 100%	[0–0.01]	Implants + ATL + Class I
5-LN	SPS, CPS	L TL electrodes	Unremarkable	Unremarkable	LTLE	−0.01649 −0.35912 0.06109	RTLE 90.53%	[0.289–0.981]	Implants + ATL + Class I

slight drop of the lowest confidence interval, but yet good average accuracy with cross-scanner data.

3.7. Using operated patients to predict non-operated patients

One of the limitations of our patient sample is that it contains data from both operated and non-operated patients. Despite the efforts of the clinical team, the identification of the seizure focus in non-operated patients remains inferential; only patients who underwent surgery with good clinical outcome can be considered to have confirmed seizure lateralization. To investigate the quality of the current classifications attributed by the clinical team, we trained the predictive algorithm on operated patients with good seizure outcome and applied the prediction to the 23 patients who did not yet receive surgery. Of the 35 patients who underwent surgery, 29 (83%) had good seizure control and were used for training (24 Engel class I and 5 Engel class II). The prediction of the 23 non-operated patients showed 100% correspondence with the current classification assigned by the clinical team. This analysis provided further evidence that, despite the inferential nature of the SOZ in non-operated patients, their clinical classification is in line with what can predicted based on available patients with good surgery outcome.

3.8. Logistic regression within MRI-positive patients

Our sample consisted in patients who were all PET-positive. Consequently, the main reason why the winning model included only PET asymmetries could be that pathologies in the other modalities were under-represented. To put both MRI and PET modalities at a comparative level, we built a predictive model only within the 35 MRI-positive patients (19 LTLE, 16 RTLE). Unthresholded asymmetries from the main dataset were introduced as potential predictors in stepwise logistic regression. Two variables entered the equation, PET-hippo and THICK-mes-Frees, reaching an average bootstrapped accuracy of 99.7 with 95% CI = [97% 100%]. This finding confirmed that MRI can be a valid predictor of seizure laterality when MRI pathology is present. However, we also applied on these patients the winning algorithm obtained from the three PET asymmetries. Average bootstrapped accuracy of this model was 99.9% with 95% CI = [100% 100%]. In sum, success rates obtained with PET asymmetries alone were comparable, if not better, than the success rates obtained with the two PET/MRI predictors above.

3.9. Clinical application: predicting new TLE patients

The applicability of the predictive algorithm was tested on a series of patients recently admitted and reviewed by the Comprehensive Epilepsy Board at Thomas Jefferson Hospital. Data from new patients followed the same processing pipeline used for the other patients in the study, though the predictive algorithm was enhanced to provide a probabilistic classification of each new patient. In short, for each new patient, 10,000 predictions were obtained with random training ratios between 52% and 97%. Each prediction was complemented with a goodness of fit (or generalizability score) of the model that produced it. This score consisted in the classification accuracy of the patients that were left out during training (i.e., accuracy of classifying 48% when the other 52% was used for training). Finally, the 10,000 predictions of the new patient were weighted by the goodness of fit of each prediction; i.e., a model that predicted 100% of the non-trained patients was counted 100 times, a model that predicted 50% of the test patients was counted 50 times. This procedure outputs the ratios of times the patient was classified as LTLE or RTLE out of the 10,000 predictions, which consists in a probabilistic classification of the patient. In addition, the confidence intervals of the regression output provide a measure of fluctuations (variability) in the probability obtained across the bootstrap. We considered both these measures when predicting new patients. The toolkit and data used in this study are available at http://dorianps.github.io/TLEprediction/.

There were seven new patients reviewed in the past months by the epilepsy board (Table 2). Their PET reports indicated three cases with left temporal hypometabolism, one bilateral hypometabolism slightly biased to the left, two right temporal hypometabolism, and one mild right temporal hypometabolism. After reviewing all the clinical information, the epilepsy board classified the first four as LTLE and the remaining three as RTLE. The prediction algorithm correctly classified all of them over 95% of the 10,000 bootstraps (range 96.33%–100%). Six of the seven patients had confidence intervals of the posterior probabilities that did not include the uncertainty threshold of 0.5, while one patient slightly crossed the threshold (95% CI = [0.44 0.99]). This patient crossed the 0.5 threshold between 2.5% and 5% of the 10,000 bootstraps, and was rated with mild hypometabolism in visual inspection reports.

3.10. Clinical application: borderline PET cases

The above new patients had lateralized hypometabolism and convergent data from other modalities, which helped the clinical team in establishing the SOZ. In other cases, the neuroimaging profile may not be clear; i.e., data may be discordant or PET visual inspections may not be informative. These are cases in which PET is rated as normal or bilaterally hypometabolic, or cases in with hypometabolism in one hemisphere and MRI pathology in the other hemisphere. To test the predictive model with such cases we searched our database for patients who had surgery with good surgery outcome, thus, providing a well-grounded confirmation of SOZ laterality. We found five cases that were initially excluded from the study because of missing DTI, and obtained the probabilistic classification from each. Their data are shown in Table 2. Patient 1-RD had discordant MRI–PET findings, with MRI revealing T2 hyperintensities in right frontal, right periventricular region, left caudate, and a mild right hippocampal atrophy. Two PET examinations were obtained from this patient, the first rated as unremarkable, the second rated with mild left temporal hypometabolism. The seizure focus was finally established through intracranial electrodes implantation, which showed the patient was LTLE. The patient underwent left anterior temporal resection and has been seizure free for 30 months. We predicted this patient as LTLE 99.5% of the time, with 95% CI of posterior probability [0.075 0.389]. Patient 2-VO was found with left MTS in MRI and was rated with bilateral temporal hypometabolism in PET. She was classified as LTLE by the clinical team and underwent laser ablation of the left hippocampus and amygdala, after which has been seizure free for 28 months. We predicted this patient as LTLE 99.9% of the time, with 95% CI of posterior probability [0.001 0.157]. Patient 3-VD had left MTS in MRI and was rated with bilateral temporal hypometabolism in PET. He was classified as LTLE by the clinical team and underwent left anterior temporal lobe resection, after which has been seizure free for 34 months. We predicted this patient as LTLE 100% of the times, with 95% CI of posterior probability [0.000 0.004]. Patient 4-CS was PET-negative and MRI-negative. She was implanted with intracranial electrodes, where seizures were found to originate in the left temporal lobe. The patient underwent left anterior temporal lobe resection and has been seizure free for 7 months. We predicted the patient as LTLE 100% of the times, with 95% CI of posterior probability [0 0.01]. Patient 5-LN was PET-negative and MRI-negative. She was implanted with intracranial electrodes, where seizures were found to originate from the left temporal lobe. She underwent left anterior temporal resection and has been seizure free for 53 months. We predicted this patient incorrectly as RTLE 90.5% of the time, but this was the only patient to show uncertain posterior probability (95% CI = [0.289 0.981]).

In sum, four of the five patients who had discordant or negative PET findings, were categorized with high certainty both in classification ratios and in the posterior probability intervals. The only patient that was miscategorized with probabilistic classification had largely uncertain posterior probabilities.

3.11. Investigation of asymmetry relationships

A comparison of average correlations between each modality pair revealed closer PET–DTI relationship in LTLE compared to RTLE ($t[30] = -3.06$, $p = .005$, 95% CI = [0.06 0.32]), and closer MRI–PET relationship in RTLE compared to LTLE ($t[46] = -3.65$, $p < .001$, 95% CI = [−0.42 − 0.12]). The MRI–DTI correlations did not differ between groups ($t[46] = -1.87$, $p = .07$). An overview of asymmetry correlations is displayed in Supplementary Fig. 1. The qualitative inspection of hierarchical clustering dendrograms showed that PET was first linked with DTI, with MRI asymmetries located on a different branch in LTLE patients. Vice versa, PET was first linked to MRI and later merged with DTI asymmetries in RTLE patients (Supplementary Fig. 1).

4. Discussion

This study investigated the predictive power of asymmetry indices from three different modalities, PET, MRI, and DTI, with the aim of building a clinically applicable predictive tool capable of distinguishing the laterality of the seizure focus in temporal lobe epilepsy. To achieve the necessary confidence required in clinical settings, we performed split-sample validations using robust statistical methods (i.e., bootstrapping), and tested the predictive model on several new patients.

4.1. Predictive power of the three modalities

Among all the asymmetries considered, PET asymmetries stood out as the most informative with respect to the lateralization of the SOZ. These asymmetries also survived multiple comparison correction, while none of the asymmetries from MRI and DTI survived the correction. The importance of PET has been reported in several previous studies. For example, reports have shown that good surgical outcome is best predicted by the presence of hypometabolism in PET images, while the presence or absence of MRI pathology adds no further value to the predictive value of PET (Carne et al., 2004; Choi et al., 2003; Feng et al., 2014; Gok et al., 2013; Kuba et al., 2011; LoPinto-Khoury et al., 2012; Struck et al., 2011; Yang et al., 2014). When comparing the predictive value of PET, MRI, and DTI, Thivard et al. (2011) found that PET carries the most useful information to identify the SOZ, while DTI added predictive value only for cases with PET-negative findings. In our study, PET had top predictive power not only in the total sample of mixed MRI-positive/MRI-negative patients, but also in the restricted group of MRI-positive patients. Thus, our data, and the above mentioned studies, highlight the importance of using metabolism as a predictor of SOZ laterality in TLE patients. Note, hypometabolism is not directly related to structural pathologies found in MRI (Lamusuo et al., 2001; O'Brien et al., 1997; Theodore et al., 2001), though recent findings suggest that even MRI-negative patients carry some structural pathology at the microscopic level (Deep et al., 2012). These findings suggest that subtle structural GM pathologies may frequently be present in TLE, and, while not visible on MRI, might still be detected at the metabolic level.

Differently from previous studies that reported success rates in the range 85%–92% when classifying TLE patients with PET data (Didelot et al., 2010; Kerr et al., 2013; Li et al., 2000; Muzik et al., 1998; Soma et al., 2012), we were able to achieve 100% success rates on the full dataset, and above 95% in split-sample validation. We took extra care in verifying this result with thousands of bootstraps and 16 split-sample steps, up to a split-half validation. While such rigorous validation has not been reported in previous studies, this is an important step for obtaining reliable and replicable classifications. Indeed, several authors have warned against the widely used leave-one-out validation, as a suboptimal method with unstable prediction error estimates (Gabrieli et al., 2015; see also http://www.russpoldrack.org/2012/12/the-perils-of-leave-one-out.html; Kohavi, 1995; Rao and Fung, 2008). Interestingly, we observed such instability as an increase in confidence interval range when the split sample approached the leave-one-out procedure (i.e., training sample = 97% in Fig. 2). This result supports the previous warnings that a reliable estimation of prediction error requires split samples larger than one or two subjects.

Among the reasons that allowed us to create a highly accurate predictive model, the analysis procedures played an important role. First, we used a new masking procedure based on smoothed cortical mantel masks, similar to probabilistic ROIs. Previous studies have used voxel-wise asymmetries (Didelot et al., 2010; Kim et al., 2003; Van Bogaert et al., 2000) or region-wise asymmetries (Muzik et al., 2005), but not asymmetries from probabilistic ROIs. The results obtained in our study suggest that probability masks of smoothed cortical mantel can be a simplistic, yet powerful, solution to automatize PET analyses. Second, we included not just asymmetries of mean, but also asymmetries of variance, skewness, and kurtosis. No previous study has used these measures to classify TLE patients yet. Importantly, results from cancer research suggest that these measures improve significantly the predictive power (Nair et al., 2012). Third, all PET asymmetries were obtained without smoothing or normalizing the data, thus, keeping the PET image as close as possible to the original scanner output. This choice seems to be advantageous for conserving subtle variance changes that can dissociate LTLE from RTLE.

PET was followed by DTI as the second best predictive modality at the individual modality level (71%), while MRI was the third best (67%). Even when multimodal asymmetries were combined, the combination of asymmetries of PET with DTI yielded higher accuracy rates than the combination of PET with MRI (lower CI 100% and 96%, respectively). This finding is congruent with the report of Thivard et al. (2011) who found DTI to be the next best predictor after PET. However, in our study MRI was an important predictor within MRI-positive patients, where mesial temporal lobe thickness was combined with hippocampal metabolism to produce a highly successful model. A similar finding was obtained by Focke et al. (2012), who studied only patients with hippocampal sclerosis, and found that T1-weighted MRI is better able to distinguish LTLE from RTLE compared to DTI. Together, our findings and the above literature indicate that MRI measures can be more sensitive than DTI in the restricted group of MRI-positive patients, while DTI can be slightly superior when all TLE patients are considered.

Attempts to predict seizure laterality from MRI or DTI have also been reported. Success rates with MRI data have been in the range 86%–100% (Duchesne et al., 2006; Focke et al., 2012; Li et al., 2000; Moser et al., 2000). Our MRI-based prediction model was less successful (67%) than the above studies, a finding that might be related to our choice of using thickness asymmetries. Previous literature has reported more widespread atrophy in LTLE than in RTLE (Bonilha et al., 2007; Keller and Roberts, 2008; Kemmotsu et al., 2011), suggesting that thickness might be a good variable for distinguishing the groups. The LTLE–RTLE difference, however, has not been consistently reported in all studies (Bernhardt et al., 2011; Lin et al., 2007). When inspecting the maps of thickness anomalies published by some of these studies (Kemmotsu et al., 2011; Lin et al., 2007) it appears that GM thinning may occur bilaterally in both LTLE and RTLE patients, reducing the power of thickness to distinguish LTLE vs. RTLE. Another factor that may limit the use of cortical thickness is the unreliable estimation of thickness in areas of MRI pathology (i.e., dysplasias). However, we expected these anomalies to be captured in asymmetries of variance, which would have been positive from a detection perspective. This benefit did not emerge from our results.

DTI is a newer method that has not been incorporated in routine clinical practice. Nevertheless, attempts have been made to use DTI data for distinguishing LTLE from RTLE. For example, Ahmadi et al. (2009) reported 90% accuracy, An et al. (2014) reported 91% accuracy, while Concha et al. (2012) reported 87% accuracy. Our success rate of 71% when using only DTI asymmetries is lower than these reports, a finding that may again be related to DTI processing steps (i.e., use of TBSS instead of regular tractography, tracts under investigation, etc.),

or the selection of variables (i.e., use of raw FA instead of FA asymmetry, use of mean diffusivity, etc.).

4.2. Inter-modality relationships

A new finding not previously reported is that LTLE patients carry a closer relationship PET–DTI, while RTLE patients carry a closer relationship PET–MRI. Previous studies have shown that the relationship of metabolism with structural integrity can break down in proximity of the epileptogenic area (Theodore et al., 2001). This indicates that the coupling of structure with metabolism is not static, and might be interrupted by the presence of pathology. On the other hand, numerous studies have shown that LTLE carry a more severe form of WM pathology than RTLE (Ahmadi et al., 2009; Lu et al., 2013; Pustina et al., 2014a; Pustina et al., 2015), and that WM is more closely associated with the clinical course (Kemmotsu et al., 2011). Therefore, one interpretation of our findings might be that the extensive WM pathology in LTLE biases the relationship of metabolism more toward the status of WM than to that of GM. Note, metabolic activity is not only related to GM structural integrity (volume, thickness, etc.), but also to functional integrity, which is based on the inputs sent and received from other brain regions. While this hypothesis is speculative, an appropriate way to investigate it in future studies is the comparison of multimodal relationships in TLE patients to the same relationships existing in the healthy population.

4.3. Different cortical thickness estimates depending on software

On the technical side, it is worth noting the differences of cortical thicknesses obtained from Freesurfer and ANTs. We found a higher concordance of thickness estimated in the lateral temporal lobe ($r = 0.673 \pm 0.08$) than the mesial temporal lobe ($r = 0.3 \pm 0.03$). The discrepancy between the two software asymmetries became evident also in the relationship of thickness with other modalities, i.e., the difference in PET–MRI relationship was largely driven by ANTs asymmetries (Supplementary Fig. 1). On the other hand, Freesurfer asymmetries entered the prediction model of the MRI-positive patients, but could not make it on the prediction models of the full sample, where ANTs asymmetries were more successful. These findings suggest that the choice of software may have a strong impact on the results, raising concerns about the stability of the findings when different methods are used in different studies.

4.4. Limitations

In considering the conclusions drawn from this study, several limitations should be noted. Our findings should not be erroneously interpreted as evidence that the other modalities besides PET are not needed. In this regard, the conclusions we draw about the accuracy rates of each modality pertain only to the specific measures used in this study (i.e., FA from DTI, thickness from MRI). There are several other measures that can be obtained from MRI and DTI, which we did not investigate (i.e., mean diffusivity, jacobian determinants, segmentation probability, etc.), not to mention recent new acquisition protocols and new analysis techniques (i.e., HARDI, spherical deconvolution of diffusion data, etc.). Moreover, we built the model on PET-positive patients, with limited and inconclusive results from PET-negative examples. A proper development of models to apply to the full range of clinical findings requires a more heterogeneous sample that includes more PET-negative and ambiguous cases.

In this study we used a relatively small sample of 58 patients, and, despite using rigorous validation methods and finding robust effects, the lack of bias cannot be entirely excluded with such small sample.

The goal of this paper was to develop a quantitative formula that would match the clinical team classification. However, a better standard for accurate prediction would be to train the model on seizure outcome. We had a limited number of patients with surgical outcome, and few of them had bad seizure outcome, providing insufficient information for training purposes.

Lastly, while we used three variables from a single modality to build our prediction model, recent studies have shown high lateralization accuracy with models based on clinical/cognitive variables (Armananzas et al., 2013), fMRI (Chiang et al., 2015), or hippocampus architectural properties (Nazem-Zadeh et al., 2014; Ver Hoef et al., 2013). Future work can build upon this knowledge to combine the predictive value of all predictive variables into a single model for routine clinical use.

4.5. Conclusions

To conclude, we investigated the separate and combined power of asymmetries from three different modalities, and found that metabolism is the single most powerful predictor of SOZ laterality. The crucial elements of PET's predictive success were the introduction of a new masking method and the inclusion of variance asymmetries. Our findings suggest that statistical processing of neuroimaging data can provide accurate classification of patients that match or exceed the detection ability of human experts, a conclusion that has been advanced by several authors (Ferrie et al., 1997; Lee et al., 2000; Matheja et al., 1998). The implementation of automated procedures of classification not only allows for an objective estimation of anomalies present in the data, but also offers a probabilistic range of confidence, which is not obtained from qualitative descriptions. Ultimately, automatic predictions can be combined with visual inspections to improve detection of subtle patterns of anomaly. In the spirit of open source science, we have made available the tools and data necessary to lateralize seizure onset in new TLE patients: http://dorianps.github.io/TLEprediction/.

Supplementary material for this article can be found online at http://dx.doi.org/10.1016/j.nicl.2015.07.010.

Acknowledgements

This work was supported, in part, by the National Institute of Neurological Disorders and Stroke (NINDS) (grant number R21 NS056071-01A1) to Dr. Joseph I. Tracy. The remaining funding was provided by other sources of internal organizational support. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

Ahmadi, M.E., Hagler Jr., D.J., McDonald, C.R., Tecoma, E.S., Iragui, V.J., Dale, A.M., Halgren, E., 2009. Side matters: diffusion tensor imaging tractography in left and right temporal lobe epilepsy. A.J.N.R. Am. J. Neuroradiol. 30 (9), 1740–1747. http://dx.doi.org/10.3174/ajnr.A165019509072.

An, J., Fang, P., Wang, W., Liu, Z., Hu, D., Qiu, S., 2014. Decreased white matter integrity in mesial temporal lobe epilepsy: a machine learning approach. Neuroreport 25 (10), 788–794. http://dx.doi.org/10.1097/WNR.0000000000000017824918460.

Armañanzas, R., Alonso-Nanclares, L., Defelipe-Oroquieta, J., Kastanauskaite, A., de Sola, R.G., Defelipe, J., Bielza, C., Larrañaga, P., 2013. Machine learning approach for the outcome prediction of temporal lobe epilepsy surgery. PLOS One 8 (4), e62819. http://dx.doi.org/10.1371/journal.pone.006281923646148.

Avants, B.B., Tustison, N.J., Song, G., Cook, P.A., Klein, A., Gee, J.C., 2011. A reproducible evaluation of ANTs similarity metric performance in brain image registration. Neuroimage 54 (3), 2033–2044. http://dx.doi.org/10.1016/j.neuroimage.2010.09.02520851191.

Behrens, T.E., Berg, H.J., Jbabdi, S., Rushworth, M.F., Woolrich, M.W., 2007. Probabilistic diffusion tractography with multiple fibre orientations: what can we gain? Neuroimage 34 (1), 144–155. http://dx.doi.org/10.1016/j.neuroimage.2006.09.01817070705.

Bernhardt, B.C., Chen, Z., He, Y., Evans, A.C., Bernasconi, N., 2011. Graph-theoretical analysis reveals disrupted small-world organization of cortical thickness correlation networks in temporal lobe epilepsy. Cereb. Cortex 21 (9), 2147–2157. http://dx.doi.org/10.1093/cercor/bhq29121330467.

Blum, D.E., Ehsan, T., Dungan, D., Karis, J.P., Fisher, R.S., 1998. Bilateral temporal hypometabolism in epilepsy. Epilepsia 39 (6), 651–659. http://dx.doi.org/10.1111/j.1528-1157.1998.tb01434.x9637608.

Bonilha, L., Keller, S.S., 2015. Quantitative MRI in refractory temporal lobe epilepsy: relationship with surgical outcomes. Quant. Imaging Med. Surg. 5 (2), 204–224. http://dx.doi.org/10.3978/j.issn.2223-4292.2015.01.0125853080.

Bonilha, L., Rorden, C., Halford, J.J., Eckert, M., Appenzeller, S., Cendes, F., Li, L.M., 2007. Asymmetrical extra-hippocampal grey matter loss related to hippocampal atrophy in patients with medial temporal lobe epilepsy. J. Neurol. Neurosurg. Psychiatry 78 (3), 286–294. http://dx.doi.org/10.1136/jnnp.2006.10399417012334.

Carne, R.P., O'Brien, T.J., Kilpatrick, C.J., MacGregor, L.R., Hicks, R.J., Murphy, M.A., Bowden, S.C., Kaye, A.H., Cook, M.J., 2004. MRI-negative PET-positive temporal lobe epilepsy: a distinct surgically remediable syndrome. Brain 127 (10), 2276–2285. http://dx.doi.org/10.1093/brain/awh25715282217.

Chang, L.C., Jones, D.K., Pierpaoli, C., 2005. RESTORE: robust estimation of tensors by outlier rejection. Magn. Reson. Med. 53 (5), 1088–1095. http://dx.doi.org/10.1002/mrm.2042615844157.

Chang, L.C., Walker, L., Pierpaoli, C., 2012. Informed RESTORE: a method for robust estimation of diffusion tensor from low redundancy datasets in the presence of physiological noise artifacts. Magn. Reson. Med. 68 (5), 1654–1663. http://dx.doi.org/10.1002/mrm.2417322287298.

Chiang, S., Levin, H.S., Haneef, Z., 2015. Computer-automated focus lateralization of temporal lobe epilepsy using fMRI. J. Magn. Reson. Imaging 41 (6), 1689–1694. http://dx.doi.org/10.1002/jmri.2469625044773.

Choi, J.Y., Kim, S.J., Hong, S.B., Seo, D.W., Hong, S.C., Kim, B.T., Kim, S.E., 2003. Extratemporal hypometabolism on FDG PET in temporal lobe epilepsy as a predictor of seizure outcome after temporal lobectomy. Eur. J. Nucl. Med. Mol. Imaging 30 (4), 581–587. http://dx.doi.org/10.1007/s00259-002-1079-812557048.

Concha, L., Kim, H., Bernasconi, A., Bernhardt, B.C., Bernasconi, N., 2012. Spatial patterns of water diffusion along white matter tracts in temporal lobe epilepsy. Neurology 79 (5), 455–462. http://dx.doi.org/10.1212/WNL.0b013e31826170b622815555.

Deep, A., Pati, S., Seif Eddeine, H., Chung, S., 2012. Neuropathological Study of Resected Cerebral Tissue from Patients with 3 Tesla MRI-Negative Refractory Epilepsy (P03.115). Neurology 78 (1), 03.115. http://dx.doi.org/10.1212/WNL.78.1_MeetingAbstracts.P03.115.

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. http://dx.doi.org/10.1016/j.neuroimage.2006.01.02116530430.

Didelot, A., Mauguière, F., Redouté, J., Bouvard, S., Lothe, A., Reilhac, A., Hammers, A., Costes, N., Ryvlin, P., 2010. Voxel-based analysis of asymmetry index maps increases the specificity of 18F−MPPF PET abnormalities for localizing the epileptogenic zone in temporal lobe epilepsies. J. Nucl. Med. 51 (11), 1732–1739. http://dx.doi.org/10.2967/jnumed.109.07093821051649.

Duchesne, S., Bernasconi, N., Bernasconi, A., Collins, D.L., 2006. MR-based neurological disease classification methodology: application to lateralization of seizure focus in temporal lobe epilepsy. Neuroimage 29 (2), 557–566. http://dx.doi.org/10.1016/j.neuroimage.2005.07.05216168675.

Engel Jr., J., 2001. Mesial temporal lobe epilepsy: what have we learned? Neuroscientist 7 (4), 340–352. http://dx.doi.org/10.1177/107385840100700410111488399.

Englot, D.J., Chang, E.F., 2014. Rates and predictors of seizure freedom in resective epilepsy surgery: an update. Neurosurg. Rev. 37 (3), 389–404. http://dx.doi.org/10.1007/s10143-014-0527-924497269.

Feng, R., Hu, J., Pan, L., Shi, J., Qiu, C., Lang, L., Gu, X., Guo, J., 2014. Surgical treatment of MRI-Negative temporal lobe epilepsy based on PET: a Retrospective Cohort Study. Stereotact. Funct. Neurosurg. 92 (6), 354–359. http://dx.doi.org/10.1159/000365575253588872.

Ferrie, C.D., Marsden, P.K., Maisey, M.N., Robinson, R.O., 1997. Visual and semiquantitative analysis of cortical FDG-PET scans in childhood epileptic encephalopathies. J. Nucl. Med. 38 (12), 1891–18949430463.

Firth, D., 1993. Bias reduction of maximum likelihood estimates. Biometrika 80 (1), 27–38. http://dx.doi.org/10.1093/biomet/80.1.27.

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. http://dx.doi.org/10.1073/pnas.20003379710984517.

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. http://dx.doi.org/10.1016/j.neuroimage.2011.07.06821835245.

Gabrieli, J.D., Ghosh, S.S., Whitfield-Gabrieli, S., 2015. Prediction as a humanitarian and pragmatic contribution from human cognitive neuroscience. Neuron 85 (1), 11–26. http://dx.doi.org/10.1016/j.neuron.2014.10.04725569345.

Ge, Y., Dudoit, S., Speed, T.P., 2003. Resampling-based multiple testing for microarray data analysis. Test 12 (1), 1–77. http://dx.doi.org/10.1007/BF02595811.

Gok, B., Jallo, G., Hayeri, R., Wahl, R., Aygun, N., 2013. The evaluation of FDG-PET imaging for epileptogenic focus localization in patients with MRI positive and MRI negative temporal lobe epilepsy. Neuroradiology 55 (5), 541–550. http://dx.doi.org/10.1007/s00234-012-1121-x23223825.

Greve, D.N., Svarer, C., Fisher, P.M., Feng, L., Hansen, A.E., Baare, W., Rosen, B., Fischl, B., Knudsen, G.M., 2014. Cortical surface-based analysis reduces bias and variance in kinetic modeling of brain PET data. Neuroimage 92, 225–236. http://dx.doi.org/10.1016/j.neuroimage.2013.12.02124361666.

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. http://dx.doi.org/10.1111/j.1528-1167.2007.01485.x18177358.

Kemmotsu, N., Girard, H.M., Bernhardt, B.C., Bonilha, L., Lin, J.J., Tecoma, E.S., Iragui, V.J., Hagler Jr., 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. http://dx.doi.org/10.1111/j.1528-1167.2011.03278.x21972957.

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, 31. http://dx.doi.org/10.3389/fneur.2013.0003123565107.

Kim, Y.K., Lee, D.S., Lee, S.K., Kim, S.K., Chung, C.K., Chang, K.H., Choi, K.Y., Chung, J.K., Lee, M.C., 2003. Differential features of metabolic abnormalities between medial and lateral temporal lobe epilepsy: quantitative analysis of (18)F-FDG PET using SPM. J. Nucl. Med. 44 (7), 1006–101212843213.

Klein, A., Ghosh, S.S., Avants, B., Yeo, B.T., Fischl, B., Ardekani, B., Gee, J.C., Mann, J.J., Parsey, R.V., 2010. Evaluation of volume-based and surface-based brain image registration methods. Neuroimage 51 (1), 214–220. http://dx.doi.org/10.1016/j.neuroimage.2010.01.09120123029.

Kohavi, R., 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. Proceedings of the 14th International Joint Conference on Artificial Intelligence — Volume 2. Morgan Kaufmann Publishers, Montreal, Quebec, Canada, pp. 1137–1143.

Kosmidis, I., 2013. brglm: Bias Reduction in Binary-Response Generalized Linear Models.

Kuba, R., Tyrlíková, I., Chrastina, J., Slaná, B., Pažourková, M., Hemza, J., Brázdil, M., Novák, Z., Hermanová, M., Rektor, I., 2011. "MRI-negative PET-positive" temporal lobe epilepsy: invasive EEG findings, histopathology, and postoperative outcomes. Epilepsy Behav. 22 (3), 537–541. http://dx.doi.org/10.1016/j.yebeh.2011.08.01921962756.

Lamusuo, S., Jutila, L., Ylinen, A., Kälviäinen, R., Mervaala, E., Haaparanta, M., Jääskeläinen, S., Partanen, K., Vapalahti, M., Rinne, J., 2001. [18F]FDG-PET reveals temporal hypometabolism in patients with temporal lobe epilepsy even when quantitative MRI and histopathological analysis show only mild hippocampal damage. Arch. Neurol. 58 (6), 933–939. http://dx.doi.org/10.1001/archneur.58.6.93311405808.

Lee, J.S., Lee, D.S., Kim, S.K., Lee, S.K., Chung, J.K., Lee, M.C., Park, K.S., 2000. Localization of epileptogenic zones in F-18 FDG brain PET of patients with temporal lobe epilepsy using artificial neural network. I.E.E.E. Trans. Med. Imaging 19 (4), 347–355. http://dx.doi.org/10.1109/42.84818510909929.

Leemans, A., Jeurissen, B., Sijbers, J., Jones, D., 2009. ExploreDTI: a graphical toolbox for processing, analyzing, and visualizing diffusion MR data. 17th Annual Meeting of Intl Soc Mag Reson Med, p. 3537.

Leiderman, D.B., Balish, M., Bromfield, E.B., Theodore, W.H., 1991. Effect of valproate on human cerebral glucose metabolism. Epilepsia 32 (3), 417–422. http://dx.doi.org/10.1111/j.1528-1157.1991.tb04671.x2044503.

Li, L.M., Caramanos, Z., Cendes, F., Andermann, F., Antel, S.B., Dubeau, F., Arnold, D.L., 2000. Lateralization of temporal lobe epilepsy (TLE) and discrimination of TLE from extra-TLE using pattern analysis of magnetic resonance spectroscopic and volumetric data. Epilepsia 41 (7), 832–842. http://dx.doi.org/10.1111/j.1528-1157.2000.tb00250.x10897154.

Lin, J.J., Salamon, N., Lee, A.D., Dutton, R.A., Geaga, J.A., Hayashi, K.M., Luders, E., Toga, A.W., Engel Jr., J., Thompson, P.M., 2007. Reduced neocortical thickness and complexity mapped in mesial temporal lobe epilepsy with hippocampal sclerosis. Cereb. Cortex 17 (9), 2007–2018. http://dx.doi.org/10.1093/cercor/bhl10917088374.

Lopez-Acevedo, M.L., Martinez-Lopez, M., Favila, R., Roldan-Valadez, E., 2012. Secondary MRI-findings, volumetric and spectroscopic measurements in mesial temporal sclerosis: a multivariate discriminant analysis. Swiss Med. Wkly. 142, w13549. http://dx.doi.org/10.4414/smw.2012.1354922688826.

LoPinto-Khoury, C., Sperling, M.R., Skidmore, C., Nei, M., Evans, J., Sharan, A., Mintzer, S., 2012. Surgical outcome in PET-positive, MRI-negative patients with temporal lobe epilepsy. Epilepsia 53 (2), 342–348. http://dx.doi.org/10.1111/j.1528-1167.2011.03359.x22192050.

Lu, J., Li, W., He, H., Feng, F., Jin, Z., Wu, L., 2013. 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 31 (1), 53–59. http://dx.doi.org/10.1016/j.mri.2012.06.03022925605.

Matheja, P., Diehl, B., Kuwert, T., Stodieck, S.R., Schäfers, M., Schäfers, K., Schuierer, G., Ringelstein, E.B., Schober, O., 1998. Measurement of temporal asymmetries of glucose consumption using linear profiles: reproducibility and comparison with visual analysis. Nuklearmedizin 37 (2), 43–489547749.

Moser, D.J., Bauer, R.M., Gilmore, R.L., Dede, D.E., Fennell, E.B., Algina, J.J., Jakus, R., Roper, S.N., Zawacki, T.M., Cohen, R.A., 2000. Electroencephalographic, volumetric, and neuropsychological indicators of seizure focus lateralization in temporal lobe epilepsy. Arch. Neurol. 57 (5), 707–712. http://dx.doi.org/10.1001/archneur.57.5.70710815137.

Muzik, O., Chugani, D.C., Shen, C., da Silva, E.A., Shah, J., Shah, A., Canady, A., Watson, C., Chugani, H.T., 1998. Objective method for localization of cortical asymmetries using positron emission tomography to aid surgical resection of epileptic foci. Comput. Aid. Surg. 3 (2), 74–82. http://dx.doi.org/10.1002/(SICI)1097-0150(1998)3:2<74::AID-IGS4>3.0.CO;2-H9784955.

Muzik, O., Pourabdollah, S., Juhasz, C., Chugani, D.C., Janisse, J., Draghici, S., 2005. Application of an objective method for localizing bilateral cortical FDG PET abnormalities to guide the resection of epileptic foci. I.E.E.E. Trans. Biomed. Eng. 52 (9), 1574–1581. http://dx.doi.org/10.1109/TBME.2005.85423316189970.

Nair, V.S., Gevaert, O., Davidzon, G., Napel, S., Graves, E.E., Hoang, C.D., Shrager, J.B., Quon, A., Rubin, D.L., Plevritis, S.K., 2012. Prognostic PET 18F-FDG uptake imaging features are associated with major oncogenomic alterations in patients with resected non-small cell lung cancer. Cancer Res. 72 (15), 3725–3734. http://dx.doi.org/10.1158/0008-5472.CAN-11-394322710433.

Nazem-Zadeh, M.R., Elisevich, K.V., Schwalb, J.M., Bagher-Ebadian, H., Mahmoudi, F., Soltanian-Zadeh, H., 2014. Lateralization of temporal lobe epilepsy by multimodal multinomial hippocampal response-driven models. J. Neurol. Sci. 347 (1–2), 107–118. http://dx.doi.org/10.1016/j.jns.2014.09.02925300772.

Ngugi, A.K., Bottomley, C., Kleinschmidt, I., Sander, J.W., Newton, C.R., 2010. Estimation of the burden of active and life-time epilepsy: a meta-analytic approach. Epilepsia 51 (5), 883–890. http://dx.doi.org/10.1111/j.1528-1167.2009.02481.x20067507.

O'Brien, T.J., Newton, M.R., Cook, M.J., Berlangieri, S.U., Kilpatrick, C., Morris, K., Berkovic, S.F., 1997. Hippocampal atrophy is not a major determinant of regional hypometabolism in temporal lobe epilepsy. Epilepsia 38 (1), 74–80. http://dx.doi.org/10.1111/j.1528-1157.1997.tb01080.x9024187.

Pollard, K.S., Dudoit, S., van der Laan, M.J., 2005. Multiple testing procedures: the multtest package and applications to genomics. Stat. Biol. Health 249–271 http://dx.doi.org/10.1007/0-387-29362-0_15.

Pustina, D., Doucet, G., Evans, J., Sharan, A., Sperling, M., Skidmore, C., Tracy, J., 2014a. Distinct types of white matter changes are observed after anterior temporal lobectomy in epilepsy. PLOS One 9 (8). http://dx.doi.org/10.1371/journal.pone.010421125089698.

Pustina, D., Doucet, G., Skidmore, C., Sperling, M., Tracy, J., 2014b. Contralateral interictal spikes are related to tapetum damage in left temporal lobe epilepsy. Epilepsia 55 (9), 1406–1414. http://dx.doi.org/10.1111/epi.1272125041176.

Pustina, D., Doucet, G., Sperling, M., Sharan, A., Tracy, J., 2015. Increased microstructural white matter correlations in left, but not right, temporal lobe epilepsy. Hum. Brain Mapp. 36 (1), 85–98. http://dx.doi.org/10.1002/hbm.2261425137314.

Rao, R.B., Fung, G., 2008. On the dangers of cross-validation. An Experimental Evaluation. SDM.. SIAM, pp. 588–596.

Soma, T., Momose, T., Takahashi, M., Koyama, K., Kawai, K., Murase, K., Ohtomo, K., 2012. Usefulness of extent analysis for statistical parametric mapping with asymmetry index using inter-ictal FGD-PET in mesial temporal lobe epilepsy. Ann. Nucl. Med. 26 (4), 319–326. http://dx.doi.org/10.1007/s12149-012-0573-822311414.

Sperling, M.R., O'Connor, M.J., Saykin, A.J., Phillips, C.A., Morrell, M.J., Bridgman, P.A., French, J.A., Gonatas, N., 1992. A noninvasive protocol for anterior temporal lobectomy. Neurology 42 (2), 416–422. http://dx.doi.org/10.1212/WNL.42.2.4161736176.

Struck, A.F., Hall, L.T., Floberg, J.M., Perlman, S.B., Dulli, D.A., 2011. Surgical decision making in temporal lobe epilepsy: a comparison of [(18)F]FDG-PET, MRI, and EEG. Epilepsy Behav. 22 (2), 293–297. http://dx.doi.org/10.1016/j.yebeh.2011.06.02221798813.

Suzuki, R., Shimodaira, H., 2006. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. Bioinformatics 22 (12), 1540–1542. http://dx.doi.org/10.1093/bioinformatics/btl11716595560.

Téllez-Zenteno, J.F., Hernández-Ronquillo, L., 2012. A review of the epidemiology of temporal lobe epilepsy. Epilepsy Res. Treat. 2012, 630853. http://dx.doi.org/10.1155/2012/63085322957234.

Theodore, W.H., 1989. SPECT and PET in epilepsy. Lancet 1 (8636), 502–503. http://dx.doi.org/10.1016/S0140-6736(89)91406-22563876.

Theodore, W.H., Gaillard, W.D., De Carli, C., Bhatia, S., Hatta, J., 2001. Hippocampal volume and glucose metabolism in temporal lobe epileptic foci. Epilepsia 42 (1), 130–132. http://dx.doi.org/10.1046/j.1528-1157.2001.080874.x11207796.

Thivard, L., Bouilleret, V., Chassoux, F., Adam, C., Dormont, D., Baulac, M., Semah, F., Dupont, S., 2011. Diffusion tensor imaging can localize the epileptogenic zone in nonlesional extra-temporal refractory epilepsies when [(18)F]FDG-PET is not contributive. Epilepsy Res. 97 (1–2), 170–182. http://dx.doi.org/10.1016/j.eplepsyres.2011.08.00521885254.

Tonini, C., Beghi, E., Berg, A.T., Bogliun, G., Giordano, L., Newton, R.W., Tetto, A., Vitelli, E., Vitezic, D., Wiebe, S., 2004. Predictors of epilepsy surgery outcome: a meta-analysis. Epilepsy Res. 62 (1), 75–87. http://dx.doi.org/10.1016/j.eplepsyres.2004.08.00615519134.

Tracy, J.I., Lippincott, C., Mahmood, T., Waldron, B., Kanauss, K., Glosser, D., Sperling, M.R., 2007. Are depression and cognitive performance related in temporal lobe epilepsy? Epilepsia 48 (12), 2327–2335. http://dx.doi.org/10.1111/j.1528-1167.2007.01254.x17697111.

Tustison, N.J., Cook, P.A., Klein, A., Song, G., Das, S.R., Duda, J.T., Kandel, B.M., van Strien, N., Stone, J.R., Gee, J.C., Avants, B.B., 2014. Large-scale evaluation of ANTs and Freesurfer cortical thickness measurements. Neuroimage 99, 166–179. http://dx.doi.org/10.1016/j.neuroimage.2014.05.044 24879923.

Van Bogaert, P., Massager, N., Tugendhaft, P., Wikler, D., Damhaut, P., Levivier, M., Brotchi, J., Goldman, S., 2000. Statistical parametric mapping of regional glucose metabolism in mesial temporal lobe epilepsy. Neuroimage 12 (2), 129–138. http://dx.doi.org/10.1006/nimg.2000.060610913319.

Ver Hoef, L.W., Williams, F.B., Kennedy, R.E., Szaflarski, J.P., Knowlton, R.C., 2013. Predictive value of hippocampal internal architecture asymmetry in temporal lobe epilepsy. Epilepsy Res. 106 (1–2), 155–163. http://dx.doi.org/10.1016/j.eplepsyres.2013.05.00823911210.

Wang, H., Suh, J.W., Das, S.R., Pluta, J., Craige, C., Yushkevich, P.A., 2012. Multi-atlas segmentation with joint label fusion. I.E.E.E. Trans. Pattern Anal. Mach. Intell. http://dx.doi.org/10.1109/TPAMI.2012.14322732662.

Westfall, P.H., Young, S.S., 1993. Resampling-Based Multiple Testing: Examples and Methods for P-Value Adjustment. Wiley.

Winston, G.P., 2015. The potential role of novel diffusion imaging techniques in the understanding and treatment of epilepsy. Quant. Imaging Med. Surg. 5 (2), 279–287. http://dx.doi.org/10.3978/j.issn.2223-4292.2015.02.0325853085.

Yang, P.F., Pei, J.S., Zhang, H.J., Lin, Q., Mei, Z., Zhong, Z.H., Tian, J., Jia, Y.Z., Chen, Z.Q., Zheng, Z.Y., 2014. Long-term epilepsy surgery outcomes in patients with PET-positive, MRI-negative temporal lobe epilepsy. Epilepsy Behav. 41, 91–97. http://dx.doi.org/10.1016/j.yebeh.2014.09.05425461196.