NeuroImage: Clinical 28 (2020) 102438

Contents lists available at ScienceDirect

NeuroImage: Clinical

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

Unsupervised machine learning reveals lesional variability in focal cortical dysplasia at mesoscopic scale

Hyo M. Lee[a], Ravnoor S. Gill[a], Fatemeh Fadaie[a], Kyoo H. Cho[b], Marie C. Guiot[c], Seok-Jun Hong[a], Neda Bernasconi[a], Andrea Bernasconi[a,*]

[a] Neuroimaging of Epilepsy Laboratory, Montreal Neurological Institute and Hospital, McGill University, Montreal, Quebec, Canada
[b] Department of Neurology, Yonsei University College of Medicine, Seoul, South Korea
[c] Department of Pathology, Montreal Neurological Institute and Hospital, McGill University, Montreal, Quebec, Canada

ARTICLE INFO

Keywords:
Epilepsy
Cortical dysplasia
MRI

ABSTRACT

Objective: Focal cortical dysplasia (FCD) is the most common epileptogenic developmental malformation and a prevalent cause of surgically amenable epilepsy. While cellular and molecular biology data suggest that FCD lesional characteristics lie along a spectrum, this notion remains to be verified *in vivo*. We tested the hypothesis that machine learning applied to MRI captures FCD lesional variability at a mesoscopic scale.

Methods: We studied 46 patients with histologically verified FCD Type II and 35 age- and sex-matched healthy controls. We applied consensus clustering, an unsupervised learning technique that identifies stable clusters based on bootstrap-aggregation, to 3 T multicontrast MRI (T1-weighted MRI and FLAIR) features of FCD normalized with respect to distributions in controls.

Results: Lesions were parcellated into four classes with distinct structural profiles variably expressed within and across patients: Class-1 with isolated white matter (WM) damage; Class-2 combining grey matter (GM) and WM alterations; Class-3 with isolated GM damage; Class-4 with GM-WM interface anomalies. Class membership was replicated in two independent datasets. Classes with GM anomalies impacted local function (resting-state fMRI derived ALFF), while those with abnormal WM affected large-scale connectivity (assessed by degree centrality). Overall, MRI classes reflected typical histopathological FCD characteristics: Class-1 was associated with severe WM gliosis and interface blurring, Class-2 with severe GM dyslamination and moderate WM gliosis, Class-3 with moderate GM gliosis, Class-4 with mild interface blurring. A detection algorithm trained on class-informed data outperformed a class-naïve paradigm.

Significance: Machine learning applied to widely available MRI contrasts uncovers FCD Type II variability at a mesoscopic scale and identifies tissue classes with distinct structural dimensions, functional and histopathological profiles. Integrating *in vivo* staging of FCD traits with automated lesion detection is likely to inform the development of novel personalized treatments.

1. Introduction

Focal cortical dysplasia (FCD) Type II is the most common epileptogenic developmental malformation and a prevalent cause of surgically amenable epilepsy. Histopathologically, FCD is typified by intracortical dyslamination and dysmorphic neurons, either in isolation (Type IIA) or together with balloon cells (Type IIB) (Blumcke et al., 2011). From a neurobiological standpoint, whether FCD II subtypes represent distinct entities or a spectrum is a matter of debate. Recent studies have shown significant cellular variability, with anomalies that may vary across lesions with the same subtype (Najm et al., 2018). Moreover, multiple subtypes may co-exist within the same FCD, with the most severe features determining the final diagnosis (Iffland and Crino, 2017). On MRI, FCD may appear as increased cortical thickness, abnormal signal intensity and blurred appearance (Bernasconi et al., 2011). The critical role of a lesion for successful surgery (Jobst and Cascino, 2015; Téllez-Zenteno et al., 2010) has motivated the development of automated methods aimed at detecting small FCD lesions often overlooked on routine radiological inspection (Gill et al., 2017; Hong et al., 2014). To date, algorithms have assumed structural homogeneity, possibly limiting sensitivity and specificity. In recent years, data-driven techniques applied to neuroimaging have offered

* Corresponding author at: Montreal Neurological Institute and Hospital, McGill University, 3801 University Street, Montreal, Quebec H3A 2B4, Canada.
E-mail address: andrea.bernasconi@mcgill.ca (A. Bernasconi).

https://doi.org/10.1016/j.nicl.2020.102438
Received 21 July 2020; Received in revised form 10 September 2020; Accepted 14 September 2020
Available online 18 September 2020

novel perspectives on brain disorders and provided diagnostic and prognostic markers in Alzheimer's disease (Zhang et al., 2016), depression (Drysdale et al., 2016) and autism (Hong et al., 2017). This is typically achieved using clustering, an unsupervised learning task that creates non-overlapping groupings based on similarity. Notably, a common challenge when using algorithms such as K-means or hierarchical clustering is the difficulty to objectively assess stability of the solutions. Conversely, consensus clustering (Bellec et al., 2010) is a more robust approach that relies on multiple iterations of the chosen clustering method on random sub-samples (or bootstraps) of the dataset. By inducing sampling variability, this approach yields metrics to assess the stability of the clusters (or classes) consistently emerging across bootstraps.

Assessing individual variability may offer a novel basis to advance our understanding of FCD neurobiology and improve lesion detection. Here, we tested the hypothesis that FCD variability is measurable at a millimetric scale. To this purpose, we applied consensus clustering on structural MRI features of FCD to identify clusters (or classes) that collectively make up a given lesion. By quantifying the *in vivo* expression of multiple pathological traits rather than assigning a given FCD to a single category, this approach combines dimensional modelling of individual lesions with categorical description of intra-lesional tissue classes. In addition, we evaluated the relationship of FCD classes to histopathology, as well as local function and large-scale connectivity as determined by resting-state fMRI. Reproducibility was assessed in two independent datasets. Finally, clinical utility was tested by comparing the performance of a detection algorithm trained on class-informed data to a class-naïve paradigm.

2. Materials and methods

2.1. Subjects

From a database of patients with drug-resistant epilepsy admitted to the Montreal Neurological Institute and Hospital between 2009 and 2018, we selected 46 consecutive individuals with histologically-proven FCD (22 females, 47.8%; mean ± SD age = 27.1 ± 8.6 years) who had research-dedicated structural and functional MRI scans, henceforth named Discovery dataset. The pre-surgical workup included seizure history, neurologic examination, neuroimaging, and video-EEG monitoring. EEG inter-ictal activity and ictal onset were concordant with the location of FCD lesions in 42 (91%) and 32 (70%) patients, respectively. We evaluated the frequency of interictal epileptic discharges (IED) on scalp EEG during multi-day monitoring (11 ± 3) and categorized them based on their frequency per recording into rare (≤25%) or frequent (≥75%). In 25 patients, surgery was preceded by invasive monitoring using stereotactic depth electrodes; all displayed high interictal activity and focal changes at seizure onset in electrodes targeting the lesion. At a mean ± SD postoperative follow-up (Engel, 2001) of 8.4 ± 2.2 years, 30 patients became seizure-free (Engel-I), 11 had rare disabling seizures (Engel-II), and 5 had worthwhile improvement (Engel-III).

Serial 5 μm paraffin-embedded histological sections of lesional tissue were stained with haematoxylin and eosin or Bielschowsky, and others immunostained using antibodies against GFAP, non-phosphorylated neurofilaments (SMI-32 monoclonal), microtubule-associated protein-2 (MAP-2), and neuronal specific nuclear protein (NeuN). FCD Type-II was defined as disrupted cortical lamination with dysmorphic neurons in isolation (IIA, n = 21) or together with balloon cells (IIB, n = 25). We evaluated severity of cortical dyslamination, blurring of cortical interface and gliosis using categorical scoring (1 = mild, 2 = moderate, 3 = severe).

In 70% of patients, routine radiological assessment was unremarkable with equal proportions between Type IIA (16/21) and Type IIB (16/25) (p = 0.37); the FCD lesion was subsequently recognized through inspection of texture maps (Bernasconi et al., 2011). There were no differences in age (27.7 ± 10.1 years *vs.* 26.6 ± 6.2 years, p = 0.64), sex (11 *vs.* 13 females, p = 0.98) and age at onset (13.6 ± 8.5 years *vs.* 11.2 ± 7.5 years, p = 0.31) between patients with FCD Type IIA and Type IIB.

The control group consisted of 35 age- and sex-matched healthy individuals (16 females, age = 28.8 ± 5.7 years). The Ethics Committee of the Montreal Neurological Institute and Hospital approved the study, and the written consent was obtained from all participants in accordance with the Declaration of Helsinki.

2.2. MRI acquisition

Images were acquired on a 3 T Siemens TimTrio scanner using a 32-channel head coil. The protocol included the following sequences: 3D T1-weighted MPRAGE (T1w; TR = 2300 ms, TE = 2.98 ms, flip angle = 9°, voxel size = 1 × 1 × 1 mm^3), 3D fluid-attenuated inversion recovery (FLAIR; TR = 5000 ms, TE = 389 ms, flip angle = 120°, 0.9 × 0.9 × 0.9 mm^3) and echo planar resting state fMRI (rsfMRI; TR = 2020 ms, TE = 30 ms, flip angle = 90°, 34 slices, voxel size = 4 × 4 × 4 mm^3, 150 volumes). For the latter, participants were instructed to lie still with their eyes closed while remaining awake. To reduce signal loss and distortions in orbitofrontal and mesiotemporal regions, slices were tilted in an oblique axial orientation.

2.3. MRI preprocessing and surface construction

T1w and FLAIR images underwent field non-uniformity correction, intensity normalization and linear registration to stereotaxic space based on the hemisphere-symmetric ICBM MNI152 template. T1w images were classified into white matter (WM), grey matter (GM) and cerebrospinal fluid (CSF) (Kim et al., 2005). FLAIR images were linearly mapped to T1w images in MNI space. The rs-fMRI was analyzed using DPARSF (rfmri.org/DPARSF); after discarding the first 5 time-points, the data underwent slice-timing and motion correction, realignment and statistical correction for nuisance effects of WM and CSF signals. To further correct for residual motion, time-points with a frame-wise displacement of > 0.5 mm were included as separate covariates (Power et al., 2012) in a linear model alongside the estimates of head motion (i.e., 3D rotations and translations, obtained from motion correction procedure) and used as final signals for the analyses. The time-points were then band-pass filtered at 0.01–0.08 Hz. Images were co-registered to the native T1w space using a boundary-based approach that maximizes alignment between intensity gradients of structural and echo-planar data (Greve and Fischl, 2009). The accuracy of multimodal registration was verified visually and corrected if needed.

We applied Constrained Laplacian Anatomic Segmentation using Proximity (CLASP) algorithm to generate models of GM-WM and GM-CSF surfaces with 41 k surface points (or vertices) per hemisphere (Kim et al., 2005). In short, CLASP iteratively expands a surface mesh to fit the GM-WM surface and subsequently estimates the GM-CSF surface by expanding the GM-WM surface along the Laplacian gradient between the two surfaces. Surface-based registration, which aligns individual participants based on cortical folding, was performed to enhance vertex-wise anatomical correspondence across participants (Lyttelton et al., 2007). Surface extraction accuracy was visually verified, and inaccuracies were manually corrected.

2.4. Surface-based feature extraction

Two experts (AB, NB) blinded to clinical information independently segmented the FCD lesions on co-registered T1w and FLAIR images; interrater Dice agreement index was 0.91 ± 0.11. Their consensus volume label (the union of the two segmentations) was intersected with cortical surfaces to generate surface-based FCD label, which served as input to the clustering algorithm. We calculated at each vertex belonging to the label morphological, intensity and functional features. To

minimize interpolation, we mapped the surfaces to the native space of each modality using the inverse transform of the initial co-registration. To enhance the signal-to-noise of the features while retaining high spatial specificity, we applied smoothing using a 2D quadratic diffusion kernel with 2 mm full-width-half-maximum. We then computed z-scores for each feature with respect to the distribution of the analogous tissues in healthy controls. For controls, we computed z-scores using a leave-one-out scheme.

To examine intracortical GM, we positioned three surfaces between the inner GM-WM and outer GM-CSF surfaces at 25%, 50%, and 75% cortical thickness, systematically sampling the axis perpendicular to the cortical ribbon. To assess the superficial WM, we generated surfaces running 1, 2 and 3 mm below the GM-WM surface guided by a Laplacian gradient between the GM-WM surface and ventricles (Liu et al., 2016). We then sampled the following vertex-wise features modelling FCD pathology *in vivo*:

a) *Cortical thickness.* To model GM thickening, we measured cortical thickness as the Euclidean distance between corresponding vertices of GM-WM and GM-CSF surfaces (Kim et al., 2015).
b) *Normalized FLAIR intensity.* Gliosis is associated with increased FLAIR signal intensity (Colombo et al., 2012). We divided FLAIR intensity by the average of GM-WM interface intensity. This value was normalized with respect to the mode of the FLAIR intensity histogram (Hong et al., 2014), corrected for CSF partial volume and mapped on each intracortical/subcortical surface. Intensities were sampled at 25, 50 and 75% intracortical and 1, 2 and 3 mm subcortical surfaces.
c) *Gradient.* To model GM-WM interface blurring, vertical gradients were computed at the GM-WM interface as T1w and FLAIR intensity differences between corresponding vertices along the 75% intracortical and 1 mm subcortical surfaces divided by the Euclidean distance between them.
d) *T1w/FLAIR ratio.* Despite histopathological evidence (Scholl et al., 2017), FCD-associated microstructural anomalies have not been previously assessed *in vivo*. To this purpose, we sampled T1w/FLAIR ratio as a proxy for myelin content (Glasser and Van Essen, 2011); decreases are interpreted as hypomyelination (Bernhardt et al., 2018). After sampling T1w/FLAIR ratio at 25, 50 and 75% intracortical and 1, 2 and 3 mm subcortical surfaces, we used a local cylindrical kernel approach to correct for outliers due to bulk blood vessels and CSF partial volumes (Glasser and Van Essen, 2011).
e) *Functional derivatives.* To assess local function, we calculated amplitude of low frequency fluctuations (ALFF) from the power spectrum of rs-fMRI timeseries as the square root of the mean amplitude of frequency components in 0.01 – 0.08 Hz range. ALFF reflects regional intensity of spontaneous neuronal activity and has been shown to relate to interictal spiking (Zhang et al., 2010). Moreover, we computed degree centrality (DC), which measures the number of connections that link a given node to the rest of the network. Hence, higher DC reflects higher connectivity of the node to the rest of the brain (Bullmore and Sporns, 2009). These features were computed voxel-wise in volume space and then mapped to the 50% intracortical surface.

2.5. Data-driven clustering of lesional vertices

After surface-based metrics were calculated from all vertices across the whole-brain, vertices located within the surface-mapped FCD lesion masks were extracted for subsequent analyses. We applied consensus clustering (Bellec et al., 2010), a procedure in which clustering is repeated across 10,000 bootstraps (i.e. random subsampling of with replacement) to estimate the stability matrix that stores the likelihood of intra-lesional FCD features to belong to the same cluster (or classes); a subsequent clustering on this matrix identifies *stable* classes that had consistently emerged across bootstraps (Fig. 1). Specifically, we first generated a data matrix where columns (n = 19,253) represent FCD vertices gathered across all patients and rows (n = 15) the corresponding groups of structural features (listed in section 2.4): FLAIR intensity (derived from 75, 50 and 25% intracortical and 1, 2 and 3 mm subcortical surfaces); T1w/FLAIR (75, 50 and 25% intracortical and 1, 2 and 3 mm subcortical), T1w and FLAIR vertical gradients (GM-WM interface), and cortical thickness.

Since groups had different number of features, a stratified duplication matched numbers across categories given by the lowest common denominator, namely six; this ensured that the feature groups' contributions to the clustering solution was not driven by differences in the number of features. Moreover, to ensure that smaller lesions (which contained lower number of vertices) contributed to the clustering result as much as larger lesions, we applied lesion-stratified bootstrapping. In this procedure, 10,000 sets of FCD vertices were extracted based on 70% random sampling with probabilities proportional to the inverse of the number of vertices based on eta^2 similarity matrices (Cohen et al., 2008). We performed spectral clustering on this similarity matrix by combining clustering solutions from the 10,000 bootstraps into a consensus matrix storing probabilities of vertices to belong to the same cluster. Finally, spectral clustering on this consensus matrix (henceforth *consensus clustering*) identified classes with distinct structural profiles that consistently emerged across bootstraps.

2.6. Evaluation of clustering solutions

For each pair of lesional vertices, we calculated the percentage of bootstrap solutions that had the same adjacency as the one in the consensus clustering solution. The percentage averaged across all pairs defined the percent agreement for the stability of fit for each K. The goodness of fit was computed using the inverse of Davies-Bouldin index, which measures the ratio between inter-cluster distance (how far clusters are separated from each other) and intra-cluster distance (how far members of a cluster are from its centroid); a higher index indicates a better fit. Bootstrap and consensus clustering were repeated for K = 2–5. We chose the K that yielded optimal percent agreement and goodness of fit.

2.7. Statistical analysis

For vertex-wise analysis, Student t-tests assessed differences in structural profiles of FCD classes with respect to analogous vertices of healthy controls. Linear mixed-effect models evaluated associations between FCD classes and function; the hemisphere harboring each patient's FCD was matched to that of a healthy control for sex and closest age (without re-using the same control hemisphere), for a total of 19,253 healthy vertices matched to 19,253 lesional vertices.

For patient-wise analysis, Student t-tests compared the relative proportion of each FCD class with histopathology. Logistic regression assessed associations between proportions of classes and clinical parameters, including age at onset, generalized seizures, interictal epileptic discharges; multinomial logistic regression was used for variables with more than two categories. Age, sex and lesion size were included as covariates. Results were corrected for multiple comparisons using the False discovery rate (FDR) at $q_{FDR} < 0.05$ (Benjamini and Hochberg, 1995).

2.8. Data-driven FCD detection

We evaluated the yield of FCD class-membership for automated lesion detection using extreme gradient boosting, a scalable tree boosting system (Chen and Guestrin, 2016). Each vertex was indexed with structural features and FCD class label. Inputs consisted of the original 15 features used for clustering in addition to their means across neighboring vertices in distance intervals of 0–2 and 2–4 mm. We implemented a two-stage classification strategy. The first was designed to

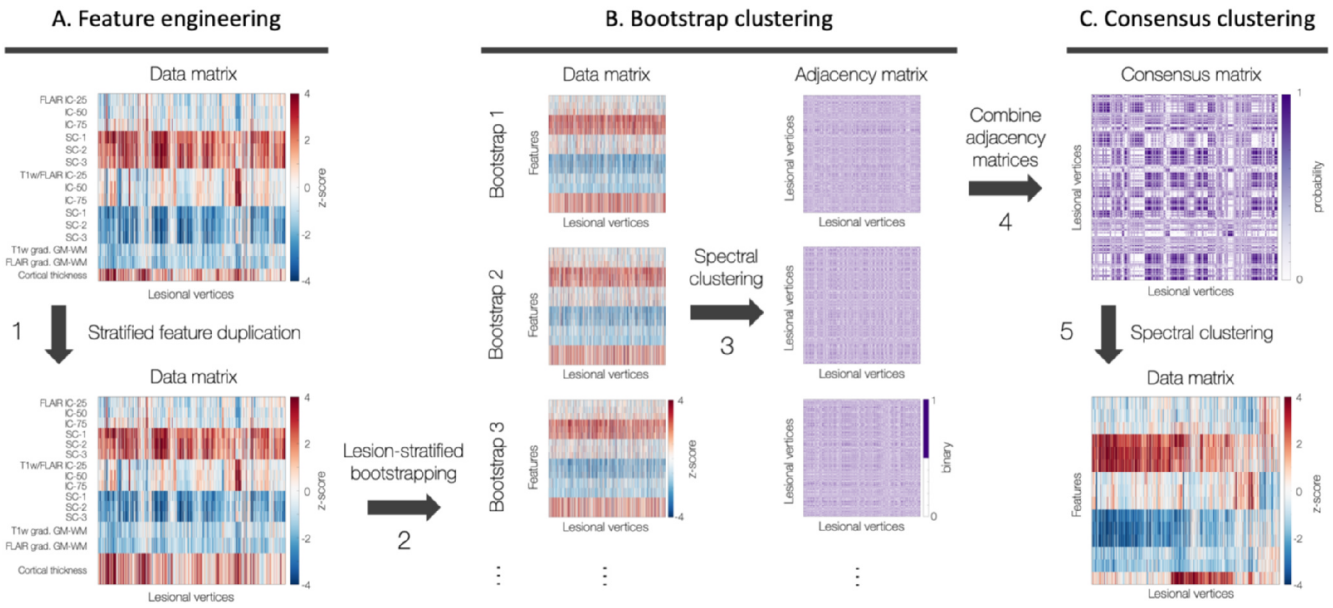

Fig. 1. Clustering framework. *A. Feature engineering.* Data matrix shows vertex-wise lesional features at intra-/subcortical levels, z-scored with respect to healthy controls (step 1, black arrow). *B. Bootstrap clustering.* Lesion-stratified bootstrapping generated 10,000 data subsets based on 70% random subsampling with replacement, ensuring equal contributions from all patients regardless of FCD size (step 2). Spectral clustering applied to each bootstrap using an eta-squared similarity matrix (step 3). *C. Consensus clustering.* Solutions from 10,000 bootstraps combined into a consensus matrix (step 4) storing probabilities of all pairs of lesional vertices to belong to the same cluster; spectral clustering on the consensus matrix identifies distinct clusters that consistently emerge across bootstraps (step 5). B and C were repeated for K = 2–5. *Abbreviations.* FLAIR: fluid-attenuated inversion recovery; GM: gray matter; int: intensity; IC/SC: intra-/subcortical; WM: white matter.

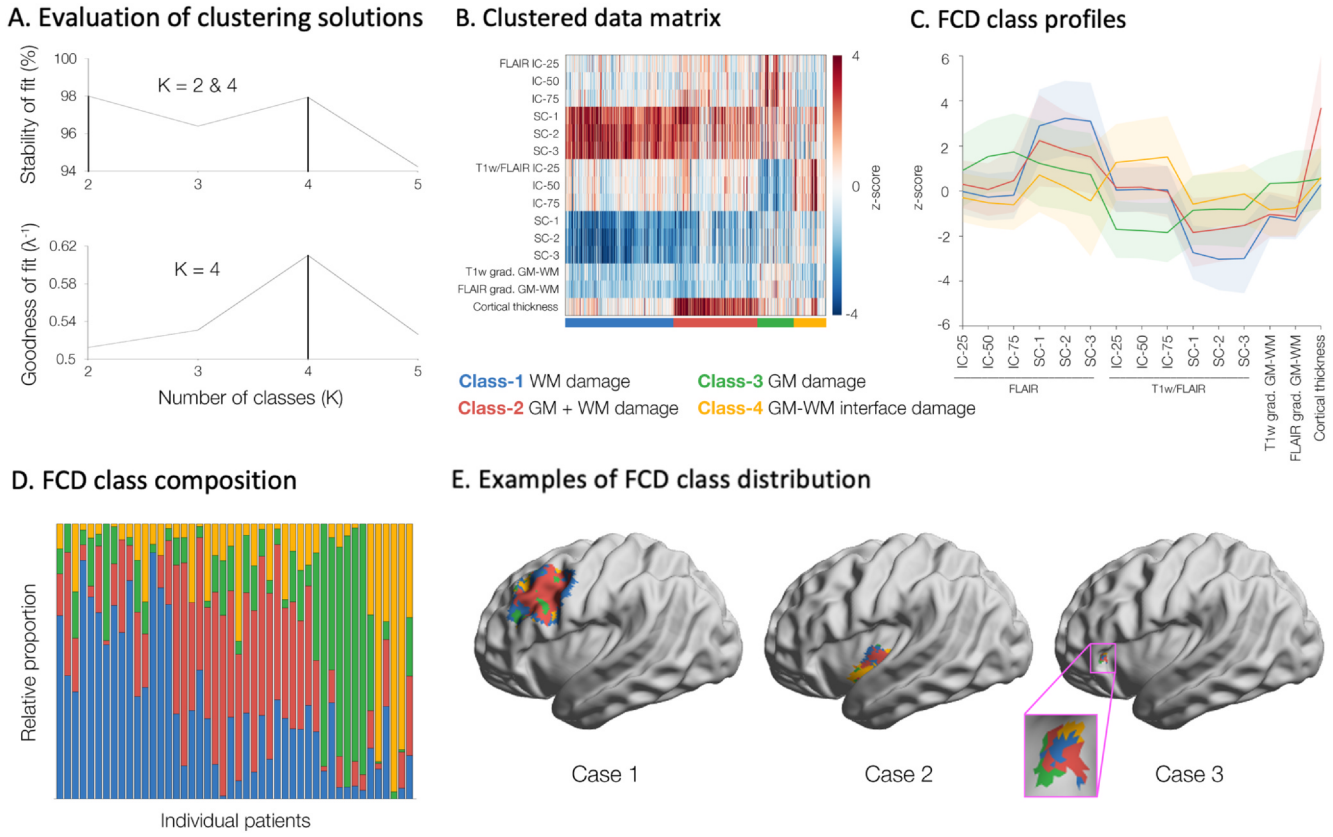

Fig. 2. Class membership. A. Stability and goodness of fit indicated K = 4 as optimal. B. The data matrix shows feature profiles of the FCD classes (color-coded); rows represent features (z-scored with respect to controls) and columns the lesional vertices. C. Line plots showing the mean and standard deviations of feature profiles of the FCD classes. D. Relative proportions of FCD classes within individual patients. E. Examples of large (case 1), medium (2) and small (3) FCD lesions.

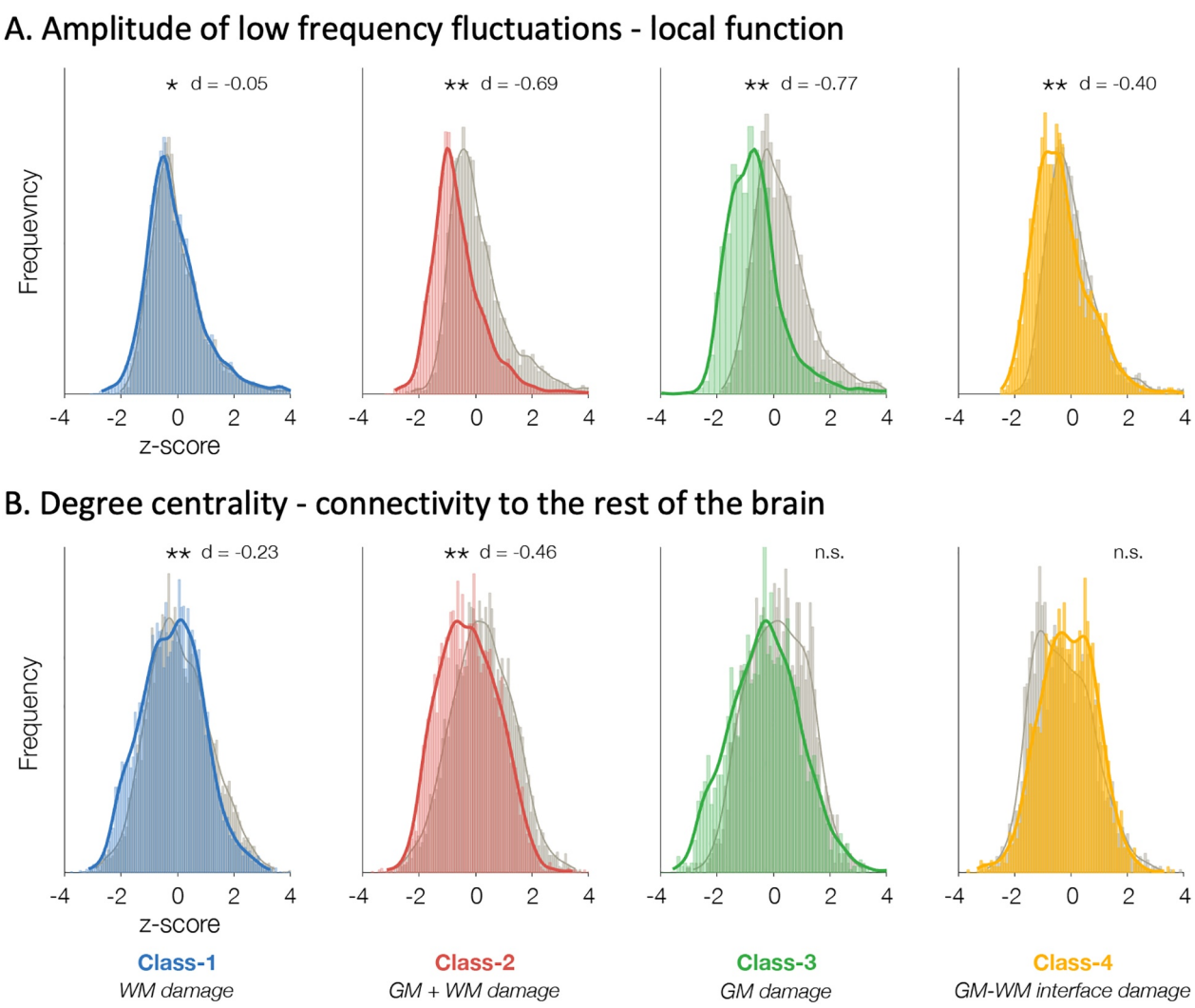

Fig. 3. Relationship of FCD classes to function. Histograms show amplitude of low frequency fluctuations indexing local function (A) and degree centrality measuring connectivity to the rest of the brain (B) for each FCD class (colored) compared to healthy controls (gray). Cohen's d effect size is indicated above each histogram. *Abbreviations*. n.s./*/**: not significant/trends ($p < 0.05$)/significant after FDR correction ($q_{FDR} < 0.05$).

maximize sensitivity and consisted of four classifiers each tuned to one of the discovered FCD classes; predictions were fed into a *meta*-classifier to produce the final prediction. The second was aimed at improving specificity by removing false positives from the first stage; it also consisted of the same four classifiers followed by a *meta*-classifier. For training, we performed random upsampling with replacement of each FCD lesion to match the number of vertices to that of the largest lesion; the same procedure was applied to the sampling of healthy vertices, thus ensuring that each patient contributed equal number of lesional and healthy vertices. In addition, to ensure that the lesional and healthy vertices contribute equal weights, we scaled the weight of the lesional class by the ratio between the total number of healthy and lesional vertices. The classifier was then trained using a 5-fold cross validation with 100 repetitions; this procedure, by which 20% of patients are classified using data from the remaining 80%, allows unbiased estimation of performance for previously unseen FCD. Finally, we compared the classification performance of the class-informed algorithm to a class-naïve classifier using two-sided McNemar's test. Student *t*-test assessed patient-wise sensitivity (percentage of detected FCD) and specificity (number of false-positive clusters).

2.9. Replication analysis

We assessed generalizability in two separate cohorts of patients with histologically-verified FCD examined at the Montreal Neurological Institute and Hospital (Replication 1; n = 14; 7 females; mean age = 24.3 ± 4.6 years) and the Severance Hospital in South Korea (Replication 2; n = 12, 4 females; mean age = 25.8 ± 8.0 years), with 3D T1w and FLAIR images acquired on a 3 T Siemens Prisma and 3 T Philips Achieva using 32-channel head coils, respectively. The MRI preprocessing and clustering procedures were identical to those applied to the Discovery cohort. To mitigate effects of scanner/site difference, imaging features underwent subject-wise z-normalization prior to z-normalization with respect to healthy controls.

3. Results

FCD Type II lesions were parcellated into four classes with distinct structural profiles, functional impact and histopathological embedding.

3.1. Data-driven FCD clustering (Fig. 2)

Clustering achieved optimal stability and goodness of fit parameters at K = 4, dichotomizing lesional vertices across patients into four

distinct classes; results were highly stable across 10,000 bootstrap instances, supporting robustness. Across classes, GM and WM profiles were significantly different from healthy controls ($q_{FDR} < 0.05$): *Class-1* with severe isolated WM anomalies, characterized by increased FLAIR intensity (indexing gliosis) and decreased T1w/FLAIR intensity (indexing hypomyelination), as well as subtle decrease in vertical gradients of T1w and FLAIR (indexing blurring), but virtually no GM changes; *Class-2* with severe GM thickening combined with moderate increase in WM FLAIR, decreased T1w/FLAIR and subtle interface blurring; *Class-3* with only moderate increase in intracortical FLAIR and decreased T1w/FLAIR, but no WM abnormalities; *Class-4* with moderate increase in intracortical T1w/FLAIR and decrease in FLAIR intensity, as well as moderate interface blurring. FCD lesions expressed at least two classes, regardless of size. Similar to the Discovery dataset, K = 4 showed optimal stability and goodness of fit in Replication 1 and 2 cohorts, with closely matching structural profiles (Figure S1).

3.2. Relationship to function (Fig. 3)

Class-1 was characterized by moderate decrease in large-scale connectivity ($q_{FDR} < 0.01$, Cohen's effect size d = -0.23), but negligible decrease in local function ($q_{FDR} < 0.001$, d = -0.05); Class-2 by severe decrease in local function ($q_{FDR} < 0.001$, d = -0.69) and moderate decrease in connectivity ($q_{FDR} < 0.01$, d = -0.46); Class-3 and Class-4 by decrease in local function, severe in the former ($q_{FDR} < 0.001$, d = -0.77) and moderate in the latter ($q_{FDR} < 0.01$, d = -0.40), but no change in connectivity.

3.3. Relationship to histopathology

Overall, MRI classes reflected typical histopathological FCD characteristics. The proportion of Class-1 vertices was more prevalent in lesions with severe WM gliosis and GM-WM interface blurring ($q_{FDR} < 0.05$), Class-2 in those with severe GM dyslamination and gliosis, as well as moderate interface blurring and WM gliosis ($q_{FDR} < 0.05$), Class-3 with moderate GM gliosis ($q_{FDR} < 0.05$), Class-4 with mild interface blurring ($q_{FDR} < 0.05$). While Class-2 was associated with Type IIB with balloon cells (log odds ratio: 5.02, $p < 0.01$), the proportion of Type IIB and IIA did not differ among the other Classes.

3.4. Relationship to clinical parameters

The relative proportions per lesion of Class-2 and Class-4 were associated with early disease onset (< 10 years; log odds ratio: 4.20 and 3.78; p = 0.02 and p = 0.04, respectively). With regards to epileptic activity, Class-1 and showed marginal association with rare compared to frequent IEDs (log odds ratio = -2.81; p = 0.07), and Class-3 with focal compared to bilateral IEDs (log odds ratio = -8.84; p = 0.07).

3.5. Data-driven FCD detection

The number of FCD vertices that the class-informed paradigm correctly predicted but class-naïve incorrectly predicted (mean ± SD = 4,770 ± 826) was higher than those the class-naïve correctly predicted but class-informed incorrectly predicted (n = 2,698 ± 172); disparity in performance was significant across all 100 repetitions (two-sided McNemar's test; p < 1e-5). At patient-level, the class-informed paradigm detected a higher number of lesions than the class-naïve (77 ± 3% *vs.* 73 ± 3%; p < 1e-5), while the number of false positive clusters did not differ (5 ± 0.3 *vs.* 5 ± 0.5).

4. Discussion

Whether FCD Type IIA and IIB represent distinct entities or a spectrum has been a matter of debate. Beside evidence for molecular variability (Crino, 2015), recent observations suggest co-expression of multiple histological subtypes within the same FCD lesion (Iffland and Crino, 2017; Najm et al., 2018); moreover, severity and arrangement of pathological features may vary between lesions assigned to the same subtype (Nakagawa et al., 2017), supporting the notion of a spectrum. Harnessing the power of bootstrap-aggregated consensus clustering, we quantified the *in vivo* expression of multiple pathological traits for a given FCD rather than assigning them to a single category, thus moving beyond previous studies assuming structural homogeneity. The high stability of clustering solutions from 10,000 bootstraps, obtained using a conservative approach based on 70% random subsampling with replacement, suggests that the FCD classes may generalize beyond the Discovery dataset of this study. Indeed, this was consolidated by the replication in two independent datasets. Lesions were parcellated into four classes with distinct structural profiles variably expressed within and across patients. Classes had differential histopathological features and functional embeddings. Clinical utility is supported by gain in performance of a lesion detection algorithm trained on class-informed data compared to a class-naïve paradigm; a main contributor resides in the explicit modeling of structural variability in the class-informed paradigm allowing FCD classes to equally contribute to the training.

Current histopathological gradings are based on descriptive criteria and often do not take into account the degree of severity of individual features. Although this approach may be sufficient in clinical practice, it may limit the neurobiological understanding of FCD. The presented MRI-based classification framework may allow us to capture a broader pathological spectrum. Indeed, the gradual structural compromise we observed across individual lesions provides the basis for a dimensional conceptualization of FCD. Phenotypical variability is further supported by the fact that discovered classes did not show consistent associations with histological subtypes. While the absence of digitized tissue samples prevented a fully quantitative comparison between MRI and histology, our imaging markers reflecting categorical variations of main FCD features emphasize the ability of post-processing to capture histopathological variations at mesoscopic scale. Indeed, MRI-derived Classes with preferential WM damage (1 and 4) were more commonly associated with histopathological features of severe GM-WM interface blurring and WM gliosis, while those with GM damage (2 and 3) displayed intracortical dyslamination and gliosis. Notably, only Class-2 typified by severe cortical thickening was associated with Type IIB, possibly in relation to increased neuronal cell diameter and balloon cells (Muhlebner et al., 2012). Notably, however, accurate histological characterization may be arduous for various reasons, including incomplete surgical sampling (Krsek et al., 2009; Oluigbo et al., 2015), difficulty of perpendicular sectioning with respect to the pial surface (Blümcke et al., 2016) due to variability in size and quality of resections, as well as the logistic burden of immunohistochemistry (Blümcke et al., 2016; Prayson, 2014).

Beside optimized performance for automated detection, the clinical relevance of FCD classes is further supported by their relation to electro-clinical parameters and age of onset. From an electrophysiological perspective, Class-1 with preferential WM damage, was associated with rare IEDs, while Class-3 with selective cortical anomalies being associated with focal discharges. Classes driven by GM anomalies had an impact on local function, whereas those with WM changes affected large-scale connectivity. A likely explanation lies in the developmental origin of FCD with stage-dependent modulation of genetically-driven molecular perturbations (Marsan and Baulac, 2018). Anomalous local function may relate to GM alterations secondary to aberrant cell proliferation during mitotic cycles, whereas WM alterations may be linked to defective later-stage neuronal migration (Guerrini and Dobyns, 2014). This is consistent with our results showing early disease onset in patients with predominant GM pathology classes. Cortical development consists of three successive but partially overlapping stages of cell proliferation, neuronal migration and cortical organization (Bystron et al., 2008). Thus, co-expression of

two or more classes within lesions suggests that molecular perturbations along overlapping stages of neurodevelopment may contribute to the structural and functional makeup of FCD.

5. Conclusions

The presented data-driven approach uncovered FCD Type II variability at a mesoscopic scale, revealing tissue classes with distinct structural dimensions, functional and histopathological profiles. From a clinical standpoint, integrating *in vivo* staging of FCD pathology with automated algorithms relying on widely available MRI contrasts is likely to pave the way for the detection of the most subtle form of cortical dysplasia characterized by isolated intra-cortical dyslamination, an elusive entity currently representing one of the main barriers to epilepsy surgery (Krsek et al., 2009; Téllez-Zenteno et al., 2010). Moreover, addressing the full spectrum of FCD traits may play a key role in establishing genotype-phenotype associations and their clinical translation, opening opportunities to inform the development of novel personalized treatments (Citraro et al., 2016) so far mainly hindered by the lack of phenotypes linked to FCD somatic variants (D'Gama and Walsh, 2018).

CRediT authorship contribution statement

Hyo M. Lee: Conceptualization, Methodology, Software, Validation, Formal analysis, Data curation, Writing - original draft. **Ravnoor Gill:** Conceptualization, Methodology, Software. **Fatemeh Fadaie:** Software. **Kyoo H. Cho:** Investigation, Data curation. **Marie C Guiot:** Investigation, Resources. **Seok-Jun Hong:** Conceptualization, Methodology. **Neda Bernasconi:** Conceptualization, Investigation, Supervision, Project administration, Funding acquisition, Writing - review & editing. **Andrea Bernasconi:** Conceptualization, Investigation, Supervision, Project administration, Funding acquisition, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was funded by the Canadian Institutes of Health Research (CIHR MOP-57840 to AB and CIHR MOP-123520 to NB), Natural Sciences and research Council (NSERC; Discovery-243141 to AB and 24779 to NB), Epilepsy Canada (Jay & Aiden Barker Breakthrough Grant in Clinical & Basic Sciences to AB), and Canada First Research Excellence Fund (HBHL-1a-5a-06 to NB). Salary supports were provided by Savoy Foundation for Epilepsy (HML, FF), Fonds de Recherche Santé Quebec (RSG), Lloyd Carr-Harris Foundation (BC), and Canadian league Against Epilepsy (SJH).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.nicl.2020.102438.

References

Bellec, P., Rosa-Neto, P., Lyttelton, O.C., Benali, H., Evans, A.C., 2010. Multi-level bootstrap analysis of stable clusters in resting-state fMRI. NeuroImage 51, 1126–1139.

Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society. Series B (Methodological) 57, 289–300.

Bernasconi, A., Bernasconi, N., Bernhardt, B.C., Schrader, D., 2011. Advances in MRI for cryptogenic epilepsies. Nature Reviews Neurology 7, 99.

Bernhardt, B.C., Fadaie, F., Vos de Wael, R., Hong, S.-J., Liu, M., Guiot, M.C., Rudko, D.A., Bernasconi, A., Bernasconi, N., 2018. Preferential susceptibility of limbic cortices to microstructural damage in temporal lobe epilepsy: A quantitative T1 mapping study. NeuroImage 182, 294–303.

Blümcke, I., Aronica, E., Miyata, H., Sarnat, H.B., Thom, M., Roessler, K., Rydenhag, B., Jehi, L., Krsek, P., Wiebe, S., Spreafico, R., 2016. International recommendation for a comprehensive neuropathologic workup of epilepsy surgery brain tissue: A consensus Task Force report from the ILAE Commission on Diagnostic Methods. Epilepsia 57, 348–358.

Blumcke, I., Thom, M., Aronica, E., Armstrong, D.D., Vinters, H.V., Palmini, A., Jacques, T.S., Avanzini, G., Barkovich, A.J., Battaglia, G., Becker, A., Cepeda, C., Cendes, F., Colombo, N., Crino, P., Cross, J.H., Delalande, O., Dubeau, F., Duncan, J., Guerrini, R., Kahane, P., Mathern, G., Najm, I., Ozkara, C., Raybaud, C., Represa, A., Roper, S.N., Salamon, N., Schulze-Bonhage, A., Tassi, L., Vezzani, A., Spreafico, R., 2011. The clinicopathologic spectrum of focal cortical dysplasias: a consensus classification proposed by an ad hoc Task Force of the ILAE Diagnostic Methods Commission. Epilepsia 52, 158–174.

Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat Rev Neurosci 10, 186–198.

Bystron, I., Blakemore, C., Rakic, P., 2008. Development of the human cerebral cortex: Boulder Committee revisited. Nat Rev Neurosci 9, 110–122.

T. Chen C. Guestrin XGBoost: A Scalable Tree Boosting System 2016 ACM, San Francisco, California, USA 785 794.

Citraro, R., Leo, A., Constanti, A., Russo, E., De Sarro, G., 2016. mTOR pathway inhibition as a new therapeutic strategy in epilepsy and epileptogenesis. Pharmacol. Res. 107, 333–343.

Cohen, A.L., Fair, D.A., Dosenbach, N.U.F., Miezin, F.M., Dierker, D., Van Essen, D.C., Schlaggar, B.L., Petersen, S.E., 2008. Defining functional areas in individual human brains using resting functional connectivity MRI. NeuroImage 41, 45–57.

Colombo, N., Tassi, L., Deleo, F., Citterio, A., Bramerio, M., Mai, R., Sartori, I., Cardinale, F., Lo Russo, G., Spreafico, R., 2012. Focal cortical dysplasia type IIa and IIb: MRI aspects in 118 cases proven by histopathology. Neuroradiology 54, 1065–1077.

Crino, P.B., 2015. The enlarging spectrum of focal cortical dysplasias. Brain 138, 1446–1448.

D'Gama, A.M., Walsh, C.A., 2018. Somatic mosaicism and neurodevelopmental disease. Nat. Neurosci. 21, 1504–1514.

Drysdale, A.T., Grosenick, L., Downar, J., Dunlop, K., Mansouri, F., Meng, Y., Fetcho, R.N., Zebley, B., Oathes, D.J., Etkin, A., Schatzberg, A.F., Sudheimer, K., Keller, J., Mayberg, H.S., Gunning, F.M., Alexopoulos, G.S., Fox, M.D., Pascual-Leone, A., Voss, H.U., Casey, B.J., Dubin, M.J., Liston, C., 2016. Resting-state connectivity biomarkers define neurophysiological subtypes of depression. Nat. Med. 23, 28.

Engel Jr., J., 2001. A proposed diagnostic scheme for people with epileptic seizures and with epilepsy: report of the ILAE Task Force on Classification and Terminology. Epilepsia 42, 796–803.

Gill, R.S., Hong, S.-J., Fadaie, F., Caldairou, B., Bernhardt, B., Bernasconi, N., Bernasconi, A., 2017. Automated Detection of Epileptogenic Cortical Malformations Using Multimodal MRI. In: Cardoso, M.J., Arbel, T., Carneiro, G., Syeda-Mahmood, T., Tavares, J.M.R.S., Moradi, M., Bradley, A., Greenspan, H., Papa, J.P., Madabhushi, A., Nascimento, J.C., Cardoso, J.S., Belagiannis, V., Lu, Z. (Eds.), Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support : Third International Workshop, DLMIA 2017, and 7th International Workshop, ML-CDS 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, September 14, Proceedings. Springer International Publishing, Cham, pp. 349–356.

Glasser, M.F., Van Essen, D.C., 2011. Mapping human cortical areas in vivo based on myelin content as revealed by T1- and T2-weighted MRI. J Neurosci 31, 11597–11616.

Greve, D.N., Fischl, B., 2009. Accurate and robust brain image alignment using boundary-based registration. NeuroImage 48, 63–72.

Guerrini, R., Dobyns, W.B., 2014. Malformations of cortical development: clinical features and genetic causes. The Lancet. Neurology 13, 710–726.

Hong, S.-J., Kim, H., Schrader, D., Bernasconi, N., Bernhardt, B., Bernasconi, A., 2014. Automated detection of cortical dysplasia type II in MRI-negative epilepsy. Neurology 83, 48–55.

Hong, S.-J., Valk, S.L., Di Martino, A., Milham, M.P., Bernhardt, B.C., 2017. Multidimensional Neuroanatomical Subtyping of Autism Spectrum Disorder. Cereb. Cortex 28, 3578–3588.

Iffland, P.H., Crino, P.B., 2017. Focal Cortical Dysplasia: Gene Mutations, Cell Signaling, and Therapeutic Implications. Annu. Rev. Pathol. 12, 547–571.

Jobst, B.C., Cascino, G.D., 2015. Resective epilepsy surgery for drug-resistant focal epilepsy: A review. JAMA 313, 285–293.

Kim, H., Caldairou, B., Hwang, J.-W., Mansi, T., Hong, S.-J., Bernasconi, N., Bernasconi, A., 2015. Accurate cortical tissue classification on MRI by modeling cortical folding patterns. Hum. Brain Mapp. 36, 3563–3574.

Kim, J.S., Singh, V., Lee, J.K., Lerch, J., Ad-Dab'bagh, Y., MacDonald, D., Lee, J.M., Kim, S.I., Evans, A.C., 2005. Automated 3-D extraction and evaluation of the inner and outer cortical surfaces using a Laplacian map and partial volume effect classification. NeuroImage 27, 210–221.

Krsek, P., Maton, B., Jayakar, P., Dean, P., Korman, B., Rey, G., Dunoyer, C., Pacheco-Jacome, E., Morrison, G., Ragheb, J., Vinters, H.V., Resnick, T., Duchowny, M., 2009. Incomplete resection of focal cortical dysplasia is the main predictor of poor post-surgical outcome. Neurology 72, 217–223.

Liu, M., Bernhardt, B.C., Hong, S.-J., Caldairou, B., Bernasconi, A., Bernasconi, N., 2016. The superficial white matter in temporal lobe epilepsy: a key link between structural and functional network disruptions. Brain 139, 2431–2440.

Lyttelton, O., Boucher, M., Robbins, S., Evans, A., 2007. An unbiased iterative group

registration template for cortical surface analysis. Neuroimage 34, 1535–1544.
Marsan, E., Baulac, S., 2018. Review: Mechanistic target of rapamycin (mTOR) pathway, focal cortical dysplasia and epilepsy. Neuropathol. Appl. Neurobiol. 44, 6–17.
Muhlebner, A., Coras, R., Kobow, K., Feucht, M., Czech, T., Stefan, H., Weigel, D., Buchfelder, M., Holthausen, H., Pieper, T., Kudernatsch, M., Blumcke, I., 2012. Neuropathologic measurements in focal cortical dysplasias: validation of the ILAE 2011 classification system and diagnostic implications for MRI. Acta Neuropathol 123, 259–272.
Najm, I.M., Sarnat, H.B., Blümcke, I., 2018. Review: The international consensus classification of Focal Cortical Dysplasia – a critical update 2018. Neuropathol. Appl. Neurobiol. 44, 18–31.
Nakagawa, J.M., Donkels, C., Fauser, S., Schulze-Bonhage, A., Prinz, M., Zentner, J., Haas, C.A., 2017. Characterization of focal cortical dysplasia with balloon cells by layer-specific markers: Evidence for differential vulnerability of interneurons. Epilepsia 58, 635–645.
Oluigbo, C.O., Wang, J., Whitehead, M.T., Magge, S., Myseros, J.S., Yaun, A., Depositario-Cabacar, D., Gaillard, W.D., Keating, R., 2015. The influence of lesion volume, perilesion resection volume, and completeness of resection on seizure outcome after resective epilepsy surgery for cortical dysplasia in children. 15, 644.
Power, J.D., Barnes, K.A., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., 2012. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. NeuroImage 59, 2142–2154.
Prayson, R.A., 2014. Classification and pathological characteristics of the cortical dysplasias. Child's Nervous System 30, 1805–1812.
Scholl, T., Mühlebner, A., Ricken, G., Gruber, V., Fabing, A., Samueli, S., Gröppel, G., Dorfer, C., Czech, T., Hainfellner, J.A., Prabowo, A.S., Reinten, R.J., Hoogendijk, L., Anink, J.J., Aronica, E., Feucht, M., 2017. Impaired oligodendroglial turnover is associated with myelin pathology in focal cortical dysplasia and tuberous sclerosis complex. Brain pathology (Zurich, Switzerland) 27, 770–780.
Téllez-Zenteno, J.F., Ronquillo, L.H., Moien-Afshari, F., Wiebe, S., 2010. Surgical outcomes in lesional and non-lesional epilepsy: A systematic review and meta-analysis. Epilepsy Res. 89, 310–318.
Zhang, X., Mormino, E.C., Sun, N., Sperling, R.A., Sabuncu, M.R., Yeo, B.T.T., the Alzheimer's Disease Neuroimaging, I, 2016. Bayesian model reveals latent atrophy factors with dissociable cognitive trajectories in Alzheimer's disease. Proceedings of the National Academy of Sciences of the United States of America 113, E6535–E6544.
Zhang, Z., Lu, G., Zhong, Y., Tan, Q., Chen, H., Liao, W., Tian, L., Li, Z., Shi, J., Liu, Y., 2010. fMRI study of mesial temporal lobe epilepsy using amplitude of low-frequency fluctuation analysis. Hum. Brain Mapp. 31, 1851–1861.