DR. PATRICK H LUCKETT (Orcid ID : 0000-0003-2262-6605)
DR. JOHN J LEE (Orcid ID : 0000-0003-2269-6267)
DR. R. EDWARD HOGAN (Orcid ID : 0000-0003-2272-5005)

Article type : Research Article

Deep learning resting state fMRI lateralization of temporal lobe epilepsy

Patrick H. Luckett[1] PhD, Luigi Maccotta[2] MD, John J. Lee[3] MD PhD, Ki Yun Park[1], Nico UF Dosenbach MD PhD[2], Beau M. Ances[2] MD PhD, R. Edward Hogan[2] MD, Joshua S. Shimony[3] MD PhD, Eric C. Leuthardt[1] MD

Author affiliations:

1 Department of Neurological Surgery, Washington University School of Medicine, St. Louis, Missouri

2 Department of Neurology, Washington University School of Medicine, St. Louis, Missouri

3 Mallinckrodt Institute of Radiology, Washington University School of Medicine, St. Louis, Missouri

Corresponding Author: Patrick H. Luckett, PhD
Department of Neurological Surgery
Campus Box 8057
660 S. Euclid Avenue
St. Louis, MO 63110

Phone: 314-747-6146
Fax: 314-362-2107
Email: luckett.patrick@wustl.edu

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/EPI.17233

ABSTRACT

Objective: Localization of focal epilepsy is critical for surgical treatment of refractory seizures. There remains a great need for non-invasive techniques to localize seizures for surgical decision-making. We investigate the use of deep learning using resting state functional MRI (RS-fMRI) to identify the hemisphere of seizure onset in temporal lobe epilepsy (TLE) patients.

Methods: 2132 healthy controls and 32 pre-operative TLE patients were studied. All participants underwent structural MRI and RS-fMRI. Healthy control data was used to generate training samples for a 3D convolutional neural network (3DCNN). RS-fMRI was synthetically altered in randomly lateralized regions in the healthy control participants. The model was then trained to classify the hemisphere containing synthetic noise. Finally, the model was tested on TLE patients to assess its performance for detecting biological seizure-onset zones, and gradient-weighted class activation mapping (Grad-CAM) identified the strongest predictive regions.

Results: The 3DCNN classified healthy control hemispheres known to contain synthetic noise with 96% accuracy, and TLE hemispheres clinically identified to be seizure onset zones with 90.6% accuracy. Grad-CAM identified a range of temporal, frontal, parietal, and subcortical regions that were strong anatomical predictors of the seizure onset zone, while the resting state networks which colocalized with Grad-CAM results included default mode, medial temporal, and dorsal attention networks. Lastly, in an analysis of a subset of patients with post-surgical outcomes, the 3DCNN leveraged a more focal set of regions to achieve classification in patients with Engel class > 1 compared to Engel class 1.

Significance: Non-invasive techniques capable of localizing the seizure-onset zone could improve pre-surgical planning in patients with intractable epilepsy. We have demonstrated the ability of deep learning to identify the correct hemisphere of the seizure onset zone in TLE patients using RS-fMRI

with high accuracy. This approach represents a novel technique of seizure lateralization that could improve preoperative surgical planning.

Key Points

- Deep learning was able to accurately classify the correct hemisphere of the seizure onset zone in TLE patients using RS-fMRI.
- The strongest predictive anatomical regions included temporal, frontal, parietal, and insular regions.
- The strongest predictive resting state networks included default mode, medial temporal, and dorsal attention networks.
- Engel class 1 showed more diffuse contributions across the brain to the lateralization score compared to Engel class>1.

Key words: Machine learning; epilepsy; resting state functional connectivity

Introduction

Epilepsy is one of the most common neurologic disorders, affecting 1.5-2.3 million people in the U.S. and 50 million people worldwide[1]. Epilepsy not fully responsive to antiepileptic medications, known as intractable epilepsy, leads to substantial morbidity, mortality, and socioeconomic burden[2]. Non-invasive methods of lateralization or localization of the seizure onset zone are crucial for diagnosis, disease stratification, and of critical interest in a preoperative workup for epilepsy surgery. Conventional non-invasive methods of localization, beyond analysis of seizure semiology, typically include long-duration scalp video-assisted electroencephalography (video-EEG), high resolution structural magnetic resonance imaging (MRI), positron emission tomography (PET), single photon

emission computed tomography (SPECT), magnetoencephalography (MEG), and neuropsychological assessment, as well as other less commonly used modalities. Consensus on seizure localization, especially for the purposes of presurgical assessment, is typically achieved after considering all available diagnostic information obtained via these techniques. However, non-invasive methods are not always concordant or able to provide definitive seizure-onset zone localization. In this situation, more invasive approaches, such as intracranial EEG, are required prior to proceeding to a definitive surgical treatment, such as a resection of ablation of the cerebral cortex. Non-invasive techniques that can localize the seizure-onset zone without the need for more invasive methods thus remain in great need and could significantly improve surgical decision-making.

Machine learning (ML) refers to a form of artificial intelligence that leverages exposure to data rather than explicit instructions to develop models. Broadly speaking, machine learning algorithms compute a mapping function that isolates and selects features of a given input and maps them to an output prediction, with the mapping function being progressively updated and improved with greater exposure to input data. In the case of deep learning, the algorithm is organized in a neural network composed of layers of nodes, with multiple layers between the input and output layers, or hidden layers, that perform feature selection and mapping automatically by updating node weights as the network is exposed to data.

ML algorithms have been applied to multiple domains, including neurophysiology of healthy controls and patients with neurologic disorders[3]. In the case of epilepsy, most applications of machine learning have focused on seizure detection, prognosis, or automated analysis and/or recognition of neuroimaging data. Some approaches, for instance, have developed classifiers for heterotopias or focal cortical dysplasias[4,5], in one instance leading to prediction of seizure freedom[6]. In a very recent study, for instance, Lariviere and colleagues used a supervised pattern learning algorithm to predict surgical outcome by analyzing resting state fMRI connectivity distance in patients with intractable focal epilepsy[7]. Gleichgerrcht et al. was also able to predict surgical outcome after epilepsy surgery using deep learning with a high positive predictive value[8]. An attractive feature of machine learning algorithms is that they can potentially extract hidden properties of a data set. In a study by Pardoe and

colleagues, a machine learning network was trained on a group of structural brain MRIs of healthy controls to estimate biological age[9]. When applied to brain MRIs of epilepsy patients the network yielded estimates that were older than the actual biological age of the patients, suggesting that the network was able to capture the premature aging effect of the disease.

There have been few attempts at using machine learning methods to lateralize or localize seizure onset. They have primarily been applied to conventional sources of non-invasive data typically available as part of a presurgical epilepsy workup, such as structural MRI, PET, or SPECT[10–13]. Resting state functional MRI (fMRI) connectivity is typically used only for localization of eloquent cortex in preoperative surgical planning in epilepsy, but has shown promise in lateralizing seizure foci in patients with intractable focal epilepsy at the group level[14]. Multiple papers have attempted to use functional connectivity to localize seizure foci[15–18], although as of now none of these methods have gained wide clinical acceptance. A technique that lateralizes the seizure-onset zone *at the individual patient level* using resting state fMRI could be of significant value to the goals of a presurgical workup. Machine learning approaches are beginning to be applied to the analysis of resting state fMRI connectivity for the purpose of seizure onset localization. **Since the literature describes decreased connectivity in patients with epilepsy as compared to controls[19], we hypothesized that by augmenting functional connectivity data in normal controls with different types of noise we could train a ML algorithm to identify the lateralization of seizure onset in patients with temporal lobe epilepsy (TLE).**

In this study, we developed a novel, deep learning approach to seizure onset lateralization that applies a three-dimensional convolutional neural network (3DCNN) to preoperative resting state functional MRI (RS-fMRI) data in patients with temporal lobe epilepsy. The neural network was trained on RS-fMRI data from healthy controls with superimposed noise and/or random permutations as ground truth. When applied to RS-fMRI data from a well-characterized set of TLE patients[14], the network achieved a preoperative seizure onset zone lateralization accuracy of 90.1%. Further, gradient-weighted class activation mapping (Grad-CAM) applied to the 3DCNN identified both structural regions and functional networks which were the strongest predictors of the seizure onset zone. Lastly,

an analysis of a subset of patients with post-surgical outcomes identified differences between Engel class 1 and Engel class > 1 outcomes.

Materials and Methods

Participants

Thirty-two patients with temporal lobe epilepsy (20 with left TLE, 12 with right TLE) were retrospectively and consecutively enrolled through the Washington University Adult Epilepsy Center. This data has been well characterized in previous studies[14]. Patients were screened for enrollment based on reported clinical semiology of their seizures and confirmed unilateral temporal lobe epilepsy by video-EEG monitoring. Clinical semiology for study inclusion included auras of epigastric rising, experiential phenomena (most commonly fear) and gustatory or olfactory sensations[20]. Patients with auras suggestive of lateral temporal onset seizures, including auditory hallucinations, visual misperceptions, or language disturbance, were excluded from the study group. Video EEG monitoring recorded all subjects having focal impaired awareness seizures. Ictal scalp EEG patterns showed focal anterior temporal distribution slowing at either EEG seizure onset, or showed progression to rhythmic focal slowing over the involved region in most seizures. Exclusion criteria included clinical or electrographic evidence of bitemporal or extratemporal seizures, developmental anomalies, cortical malformations or other focal lesion on structural MRI, age < 18 years, inability to perform MRI, suspected pregnancy, history of substance or alcohol abuse, and non-proficiency in the English language. A subset of the patients had MRI evidence of medial temporal sclerosis (MTS) and/or hippocampal atrophy assessed via MRI image properties. No patients showed bilateral pathologic changes.

RS-fMRI data from healthy control participants (N=2132) was acquired through ongoing research studies conducted at WUSTL, as well as the Brain Genomics Superstruct Project[21]. Normative cognitive status for the control participants was confirmed by a structured clinical rating scale (i.e.,

the Clinical Dementia Rating Scale[22]) or online cognitive test protocol[21]. The appropriate Institutional Review Boards approved all studies, and all participants provided written informed consent.

MRI Acquisition and Processing

Images from TLE participants were acquired with a Siemens MAGNETOM Trio 3 T scanner (Erlangen, Germany). A high-resolution (0.42 × 0.42 × 0.9 mm) T1-weighted magnetization-prepared rapid gradient echo (MPRAGE) (TI 800 ms, TE 3.29 ms) structural scan was obtained for each participant for the purpose of anatomic segmentation, registration to the functional images, and atlas transformation. Two BOLD resting state functional scans (echoplanar, TR 2200 ms, TE 27 ms, flip angle 90 deg, 4x4x4 mm voxels) were also acquired for each participant. Each of the two functional runs was comprised of 164 volumes (approximately 6 minutes in duration). Study participants were asked to relax while fixating on a cross hair. fMRI BOLD data was pre-processed in several steps using standard methods[23]. Motion was corrected with rigid-body realignment. Spurious variance was minimized by removal of the linear trend, temporal low-pass filtering, spatial smoothing, and regression of nuisance parameters (head-motion, white matter, ventricular and global signals) and their temporal derivatives[24].

Imaging from healthy controls was performed on 3T Siemens MR Scanners (models varied between studies) equipped with a standard 12-channel head coil. A high-resolution, three-dimensional, sagittal, T1-weighted, magnetization-prepared rapid gradient echo (MPRAGE) scan was acquired (echo time [TE] = 16 ms, repetition time [TR] = 2,400 ms, inversion time = 1,000 ms, flip angle = 8°, 256 × 256 acquisition matrix, 1 mm^3 voxels). RS-fMRI scans were collected using an echo planar sequence (voxel size = 3-4 mm^3, TR = 2200-3000 ms, FA = 80°-90°). Each participant had approximately 14 min of RS-fMRI data. Pre-processing methods was identical to those described above.

Analysis

A three-dimensional convolutional neural network (3DCNN) with a densely connected architecture[25] was used for our analysis and was trained to classify which hemisphere contained the seizure onset zone. The 3DCNN consisted of 72 layers and 55,680,450 learnable parameters. Layers within the

network included 3x3x3 and 1x1x1 convolutions (stride 1), batch normalization, swish activation, 2x2x2 max pooling (stride 2), concatenations, and dropout (50%). Input to the model was a 48x64x48x180 BOLD time series (height, width, depth, time). 200,000 training instances were generated for each hemisphere. A minibatch size of four was used during training, and minibatches were shuffled after every full epoch. The model learn rate was initialized at .01 and was reduced by 50% at every epoch. Twenty percent of the generated training data was reserved for validation. Validation occurred at the completion of every epoch, and training terminated if the accuracy did not improve after three validations. The output of the model represented the probability of the training sample having a seizure onset zone located in each hemisphere of the brain. All analysis and models were implemented in Matlab R2020b (www.mathworks.com).

Training data was generated by augmenting data from healthy control participants. Specifically, to generate any single instance of training data, a single seed voxel was selected at random in the brain. Then, 3 integer values ranging from 0 to 3 were generated at random that would represent the height, width, and depth of the 3-dimensional region of the simulated seizure onset zone in the brain centered around the seed voxel (i.e. if the random integers chosen were [2, 1, 3], the dimension of the seizure onset zone would be 5x3x7). Next, any voxel contained within the simulated seizure onset zone would have its time series modified by various means to create anomalous data in that region of the brain. Numerous methods were uniformly used to modify the time series, including adding 1) normally distributed noise, 2) uniformly distributed noise, 3) power-law noise, 4) Brownian noise, 5) noise generated by a Levy walk process, 6) randomly permuting the time series, 7) replacing the time series with cosine waves with various periods, or 8) replacing the whole time series with constant values. **The choice of using multiple noise models is due to the fact that 1) using a single noise model would lead to the 3DCNN being overfit to that specific noise model, and 2) we do not know what an epileptic "signature" looks like in BOLD data, and are therefore unable to choose a single noise model which is most reflective of epilepsy (assuming such a noise model exist). Therefore, the use of multiple noise models leads to a nonspecific anomaly detector.** Each voxel within the simulated seizure onset zone was modified independently, and every voxel in a sample was perturbed with the same noise model. This process was repeated for the healthy control data until

there were 200,000 training samples in each hemisphere. Notably, only modified heathy control data was used to train the 3DCNN. After training was complete, the model was then tested on TLE data. Figure 1 shows an example of noise injection on BOLD data from a healthy control participant.

Features used by the 3DCNN to appropriately classify TLE data were identified using gradient-weighted class activation mapping (Grad-CAM)[26]. Grad-CAM determines the important features in data by evaluating the gradients of the target within the 3DCNN. Specifically, Grad-CAM computes the gradient with respect to the class score and the convolutional features of a specified layer. After spatially pooling the gradients, the weights are used to combine the activation maps and determine which features are most relevant to the classification goal. The final concatenation layer of the 3DCNN was used to extract features with Grad-CAM. Voxelwise Grad-CAM weights were averaged based on Freesurfer[27] defined ROIs to identify the anatomical regions that were typically of greatest importance for classification. Grad-CAM weights were also averaged across predefined ROIs[28] to identify which resting state networks (RSNs) contributed most to the classification. The RSNs evaluated included dorsal somatomotor (SMD), ventral somatomotor (SMV), cinguloopercular (CON), auditory (AUD), default mode (DMN), parietal memory (PMN), visual (VIS), frontoparietal (FPN), salience (SAL), ventral attention (VAN), dorsal attention (DAN), medial temporal (MTL), basal ganglia (BGA), and thalamus (THA). **Averaging Grad-CAM weights based on both anatomical and RSN parcellations was performed in order to summarize the voxelwise results produced by Grad-CAM, and place our results in the context of past research.**

Results

Participant demographics

Demographics for TLE and healthy control participants are presented in table 1.

Model results

The 3DCNN classified healthy control hemispheres known to contain synthetic noise with 96% accuracy. These results required 17 training epochs. Once training was complete, the model was then

tested on TLE data. These results can be seen in Supplemental Figure 1. The overall accuracy of the model on TLE patients was 90.6% (90% accuracy classifying left TLE and 91.7% accuracy classifying right TLE).

Figure 2 shows the mean feature maps identified with Grad-CAM for left and right TLE. This figure demonstrates that the areas used for identification of the left versus right TLE are, as expected, more weighted towards the left versus right side respectively. However, identification of right TLE (RTLE) is more focused on information from the right temporal lobe as compared with the left TLE (LTLE) which is more diffusely spread beyond just the left temporal lobe, involving areas of the left > right frontal lobes (see Supplemental Figure 2 and 3). The asymmetry is also illustrated in Figure 3, which shows the average Grad-Cam weights based on anatomical parcellations. In RTLE, the strongest features were primarily localized to the middle and inferior temporal lobe. Comparatively, Grad-CAM weights in LTLE were less localized and covered more regions of the brain, including temporal, frontal, and subcortical (anterior cingulate, insula, entorhinal cortex) regions.. Asymmetry between the two sides is also seen in Figure 4, which shows the average Grad-CAM weights for individual resting state networks based on predefined ROIs. Overall, the highest average weights were seen in the default mode network, medial temporal network, and dorsal attention network. Lastly, a subset of LTLE patients with different post-surgical Engel outcomes were compared. Grad-CAM feature maps for four patients with Engel class 1 were averaged, and maps from 4 patients with Engel class >1 were averaged. Figure 5 depicts the results. The mean Grad-CAM map for Engel class > 1, i.e. with poorer postsurgical outcome, shows greater focality when compared to the more diffuse extent seen in Engel class 1 patients.

Discussion

In this study, we used a novel data augmentation methodology and a 3DCNN to classify the hemisphere containing the seizure onset zone in preoperative TLE patients with high accuracy (Supplemental Figure 1). This degree of accuracy is substantial and if confirmed in larger studies,

could be sufficient for clinical use and could potentially improve treatment outcomes for patients with TLE. Further, features extracted from the 3DCNN were used to identify both anatomical regions and RSNs which contributed most to the classification. At the region level, temporal, frontal, parietal, and insular regions were identified as the strongest predictors, while at the level of resting state networks strongest predictors of lateralization included default mode, medial temporal, and dorsal attention networks (Figures 3 and 4). Lastly, in an analysis of a subset of patients with post-surgical outcomes, our model indicates that patients with poorer outcomes (Engel class >1) had a more focal set of left regions affecting classification compared to patients with Engel class 1 outcomes (Figure 5), for whom classification leveraged a diffuse area across both hemispheres.

Our analysis revealed notable differences when comparing left and right TLE (Figures 2-4). In LTLE, mean Grad-CAM maps showed strong contributions to the lateralization by numerous regions, including the medial occipitotemporal gyrus, frontal lobe, temporal lobe, and anterior cingulate cortex (ACC). In contrast, strong activations in RTLE were more localized, and primarily encompassed the temporal lobe and some parietal regions. The asymmetries seen in Figure 2 are also reflected in the numerical results presented in Figure 3. The more diffuse spread of Grad-CAM weights in the LTLE is clearly seen in comparison with the values in the RTLE patients, which are more focused in the right temporal lobe area. Past studies of TLE using multiple imaging modalities have revealed similar imbalances with respect to the effect of LTLE and RTLE in the brain[29–32]. Besson et al.[33] compared LTLE, RTLE, and controls using diffusion weighted imaging. Their results showed a significant reduction in connectivity between LTLE patients and controls, while loss of connectivity in RTLE patients was less pronounced and restricted to fewer regions. Both these and our findings may at least in part reflect the commonly observed difference in semiology between temporal seizures that involve the dominant vs. non-dominant hemisphere, with the former typically leading to much greater ictal disruption of awareness and cognitive function[34]. The 3D-CNN analysis may thus have revealed a group-specific difference in extent of the epileptogenic network reflected in a typically observed clinical phenotype.

Similarly, differences between LTLE and RTLE were observed with regards to the contribution of different resting state networks to classification. In general, LTLE showed stronger contributions by most networks (Figure 4). Again, differences in left and RTLE at the resting state network level have also been observed in other work[35]. Haneef et al.[30] showed LTLE had greater alterations in hippocampal networks when compared to RTLE. Their study also found no significant changes in frontal connectivity in RTLE patients. As seen in Figure 2, some of the largest differences seen in the Grad-CAM feature maps between LTLE and RTLE were observed in frontal regions, with orbitofrontal cortex contributing significantly to lateralization in LTLE patients but not in RTLE. These results are consistent with past research, which suggests that LTLE leads to more prominent frontal lobe dysfunction[30]. Haneef et al.[36] also showed the DMN was more effected in LTLE compared to RTLE. In all, our results may reflect the commonly observed finding that LTLE appears to cause more extensive network impairment than RTLE, which may be of relevance for preoperative planning.

The primary resting state networks associated with greater Grad-CAM feature weights (i.e. that were of greater impact on classification) included the DMN, DAN, and MTL networks in addition to subcortical regions, such as the BGA. The DMN spans a large portion of the brain and is most active when an individual is not engaged in goal-directed tasks[28]. The DMN has been implicated in numerous neurological disorders[37], including epilepsy[38,39]. The literature is consistent in reporting decreases in DMN connectivity in patients with epilepsy[19,40], although contralateral increase have been report[41]. In TLE, alterations in the DMN have been associated with greater postsurgical memory decline[42]. However, the literature is inconsistent if changes in the DMN due to TLE are dependent[43] or independent of duration of illness[44]. In direct opposition to the DMN, the DAN is most associated with task positive brain activity (primarily top-down, goal-directed attention processes)[28]. Studies have shown that TLE decreases functional connectivity in the DAN, and as a result, leads to poorer performance on cognitive test associated with top-down attention processes[45,46]. The MTL and basal ganglia encompass deep gray structures, including regions encompassing the hippocampus, parahippocampal regions, and entorhinal cortex[28], all regions that are either typically or often adversely affected by TLE. The disease has thus also been shown to be associated with alterations in

functional connectivity in regions spanned by the MTL network[30,47–50]. In our study, the 3DCNN identified functional and structural regions of the brain known to be affected by TLE. The fact that the model was trained only on health control data augmented with synthetic noise and was still able to identify regions known to associate with TLE suggests that TLE causes chronic perturbations in brain dynamics that are critically relevant to lateralization/localization of the seizure onset zone.

Figure 5 depicts the presurgical mean Grad-CAM maps for LTLE patients with Engel class 1 versus Engel class > 1 postsurgical outcome. The map for Engel class 1 showed more diffuse but weaker contribution across the brain to the lateralization score when compared to Engel class > 1, which is more focused on the left temporal and frontal lobes with relatively higher contribution levels. A potential interpretation of these results is that patients who do better postoperatively have a more diffuse set of regions with mild abnormal connectivity, and the 3D-CNN is still able to detect them and leverage them for the purpose of lateralization. If confirmed in a larger data set, the spatial degree and magnitude of features identified by the 3DCNN could serve as a biomarker to aid in surgical planning. An alternative hypothesis would posit that the group of patients with poorer postsurgical outcome had more intense focal connectivity changes that spanned a more local set of regions, which in turn contributed more on average to the classification score.

The main limitation of our study is the small size of the test set and the number of patients with postsurgical outcomes. **A possible consequence is potential asymmetry between the left and right TLE groups (with lower MTS and poorer Engel class outcomes in the left TLE group).** Further work with larger data sets will be required for validation of this technique prior to clinical applications. Since the training method is general, and not specific to TLE, future work could involve designing models which are not specific to TLE and can generalize to multiple other forms of epilepsy. Finally, we hypothesize that models capable of segmenting individual regions, rather than just hemisphere classification, could be trained to facilitate fine-grained localization of seizure onset zones, of potential use for individually tailored presurgical planning.

Conclusions

Non-invasive techniques capable of localizing the seizure-onset zone could significantly improve the evaluation of patients with intractable temporal lobe epilepsy. In this work, we have demonstrated the ability of deep learning to identify the correct hemisphere of the seizure onset zone in TLE patients using RS-fMRI. The model was able to make accurate predictions after training exclusively on healthy control data augmented with synthetic noise. This data-driven approach represents a novel noninvasive technique for seizure lateralization in TLE that could improve preoperative surgical planning and patient outcomes.

Acknowledgements

Research reported in this publication was supported by the National Institutes of Health under Award Numbers P50 HD103525, P01 AG003991, P01 AG026276, R01 CA203861, R01 DA054009, R01 AG057680, R01 MH118031, R01 NR015738, R01 NR014449. Further, we acknowledge the Paula and Rodger O. Riney fund and the Daniel J. Brennan fund.

Authorship statement

All authors contributed equally to this research. PHL, ECL, JSS, and JJL conceptualized and developed the 3DCNN and noise injection approach. LM, KYP, NUFD, BMA, and REH assisted in model evaluation and interpretation.

Conflict of Interest

The authors report no conflicts of interest. We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines.

References

1. Ivanova JI, Birnbaum HG, Kidolezi Y, et al. Economic burden of epilepsy among the privately insured in the US. *Pharmacoeconomics*. 2010;28(8):675-685.
2. Begley CE, Famulari M, Annegers JF, et al. The cost of epilepsy in the United States: an estimate from population-based clinical and survey data. *Epilepsia*. 2000;41(3):342-351.
3. Rajkomar A, Dean J, Kohane I. Machine Learning in Medicine. *N Engl J Med*. 2019.

4. El Azami M, Hammers A, Jung J, et al. Detection of lesions underlying intractable epilepsy on T1-weighted MRI as an outlier detection problem. *PLoS One*. 2016;11(9):e0161498.
5. Jin B, Krishnan B, Adler S, et al. Automated detection of focal cortical dysplasia type II with surface-based magnetic resonance imaging postprocessing and machine learning. *Epilepsia*. 2018;59(5):982-992.
6. Hong S-J, Bernhardt BC, Schrader DS, et al. Whole-brain MRI phenotyping in dysplasia-related frontal lobe epilepsy. *Neurology*. 2016;86(7):643-650.
7. Larivière S, Weng Y, Vos de Wael R, et al. Functional connectome contractions in temporal lobe epilepsy: Microstructural underpinnings and predictors of surgical outcome. *Epilepsia*. 2020;61(6):1221-1233.
8. Gleichgerrcht E, Munsell B, Bhatia S, et al. Deep learning applied to whole-brain connectome to determine seizure control after epilepsy surgery. *Epilepsia*. 2018;59(9):1643-1654.
9. Pardoe HR, Cole JH, Blackmon K, et al. Structural brain changes in medically refractory focal epilepsy resemble premature brain aging. *Epilepsy Res*. 2017;133:28-32.
10. Rudie JD, Colby JB, Salamon N. Machine learning classification of mesial temporal sclerosis in epilepsy patients. *Epilepsy Res*. 2015;117:63-69.
11. Keihaninejad S, Heckemann RA, Gousias IS, et al. Classification and lateralization of temporal lobe epilepsies with and without hippocampal atrophy based on whole-brain automatic MRI segmentation. *PLoS One*. 2012;7(4):e33096.
12. Kerr WT, Nguyen ST, Cho AY, et al. Computer-aided diagnosis and localization of lateralized temporal lobe epilepsy using interictal FDG-PET. *Front Neurol*. 2013;4:31.
13. Lopes R, Steinling M, Szurhaj W, et al. Fractal features for localization of temporal lobe epileptic foci using SPECT imaging. *Comput Biol Med*. 2010;40(5):469-477.
14. Maccotta L, He BJ, Snyder AZ, et al. Impaired and facilitated functional networks in temporal lobe epilepsy. *NeuroImage Clin*. 2013;2:862-872.
15. Boerwinkle VL, Mirea L, Gaillard WD, et al. Resting-state functional MRI connectivity impact on epilepsy surgery plan and surgical candidacy: prospective clinical work. *J Neurosurg Pediatr*. 2020;25(6):574-581.
16. Stufflebeam SM, Liu H, Sepulcre J, et al. Localization of focal epileptic discharges using

functional connectivity magnetic resonance imaging. *J Neurosurg*. 2011;114(6):1693-1697.
17. Weaver KE, Chaovalitwongse WA, Novotny EJ, et al. Local functional connectivity as a pre-surgical tool for seizure focus identification in non-lesion, focal epilepsy. *Front Neurol*. 2013;4:43.
18. Bettus G, Bartolomei F, Confort-Gouny S, et al. Role of resting state functional connectivity MRI in presurgical investigation of mesial temporal lobe epilepsy. *J Neurol Neurosurg Psychiatry*. 2010;81(10):1147-1154.
19. Voets NL, Beckmann CF, Cole DM, et al. Structural substrates for resting network disruption in temporal lobe epilepsy. *Brain*. 2012;135(8):2350-2357.
20. Hogan RE. Mesial temporal sclerosis: clinicopathological correlations. *Arch Neurol*. 2001;58(9):1484-1486.
21. Buckner RL, Roffman JL, Smoller JW. Brain Genomics Superstruct Project (GSP). Harvard Dataverse.
22. Morris JC. The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology*. 2012;41(1):1588-1592.
23. SNYDER AZ. Difference Image vs Ratio Image Error Function Forms in PET—PET Realignment. In: *Quantification of Brain Function Using PET*. Elsevier; 1996:131-137.
24. Fox MD, Zhang D, Snyder AZ, et al. The global signal and observed anticorrelated resting state brain networks. *J Neurophysiol*. 2009;101(6):3270-3283.
25. Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks. In: *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*. ; 2017.
26. Selvaraju RR, Cogswell M, Das A, et al. Grad-cam: Visual explanations from deep networks via gradient-based localization. In: *Proceedings of the IEEE International Conference on Computer Vision*. ; 2017:618-626.
27. Fischl B. FreeSurfer. *Neuroimage*. 2012;62(2):774-781.
28. Seitzman BA, Snyder AZ, Leuthardt EC, et al. The State of Resting State Networks. *Top Magn Reson Imaging*. 2019.
29. Kemmotsu N, Girard HM, Bernhardt BC, et al. MRI analysis in temporal lobe epilepsy:

cortical thinning and white matter disruptions are related to side of seizure onset. *Epilepsia*. 2011;52(12):2257-2266.

30. Haneef Z, Lenartowicz A, Yeh HJ, et al. Functional connectivity of hippocampal networks in temporal lobe epilepsy. *Epilepsia*. 2014;55(1):137-145.
31. Keller SS, Wieshmann UC, Mackay CE, et al. Voxel based morphometry of grey matter abnormalities in patients with medically intractable temporal lobe epilepsy: effects of side of seizure onset and epilepsy duration. *J Neurol Neurosurg Psychiatry*. 2002;73(6):648-655.
32. Riederer F, Lanzenberger R, Kaya M, et al. Network atrophy in temporal lobe epilepsy: a voxel-based morphometry study. *Neurology*. 2008;71(6):419-425.
33. Besson P, Dinkelacker V, Valabregue R, et al. Structural connectivity differences in left and right temporal lobe epilepsy. *Neuroimage*. 2014;100:135-144.
34. Kerling F, Mueller S, Pauli E, et al. When do patients forget their seizures? An electroclinical study. *Epilepsy Behav*. 2006;9(2):281-285.
35. Tracy JI, Doucet GE. Resting-state functional connectivity in epilepsy: growing relevance for clinical decision making. *Curr Opin Neurol*. 2015;28(2):158-165.
36. Haneef Z, Lenartowicz A, Yeh HJ, et al. Effect of lateralized temporal lobe epilepsy on the default mode network. *Epilepsy Behav*. 2012;25(3):350-357.
37. Mohan A, Roberto AJ, Mohan A, et al. Focus: the aging brain: the significance of the default mode network (DMN) in neurological and neuropsychiatric disorders: a review. *Yale J Biol Med*. 2016;89(1):49.
38. Gonen OM, Kwan P, O'Brien TJ, et al. Resting-state functional MRI of the default mode network in epilepsy. *Epilepsy Behav*. 2020;111:107308.
39. Parsons N, Bowden SC, Vogrin S, et al. Default mode network dysfunction in idiopathic generalised epilepsy. *Epilepsy Res*. 2020;159:106254.
40. Maccotta L, Lopez MA, Adeyemo B, et al. Postoperative seizure freedom does not normalize altered connectivity in temporal lobe epilepsy. *Epilepsia*. 2017;58(11):1842-1851.
41. Bettus G, Guedj E, Joyeux F, et al. Decreased basal fMRI functional connectivity in epileptogenic networks and contralateral compensatory mechanisms. *Hum Brain Mapp*. 2009;30(5):1580-1591.

42. McCormick C, Quraan M, Cohn M, et al. Default mode network connectivity indicates episodic memory capacity in mesial temporal lobe epilepsy. *Epilepsia*. 2013;54(5):809-818.
43. Kay BP, DiFrancesco MW, Privitera MD, et al. Reduced default mode network connectivity in treatment-resistant idiopathic generalized epilepsy. *Epilepsia*. 2013;54(3):461-470.
44. James GA, Tripathi SP, Ojemann JG, et al. Diminished default mode network recruitment of the hippocampus and parahippocampus in temporal lobe epilepsy. *J Neurosurg*. 2013;119(2):288-300.
45. Zhang Z, Lu G, Zhong Y, et al. Impaired attention network in temporal lobe epilepsy: a resting FMRI study. *Neurosci Lett*. 2009;458(3):97-101.
46. Liao W, Zhang Z, Pan Z, et al. Altered functional connectivity and small-world in mesial temporal lobe epilepsy. *PLoS One*. 2010;5(1):e8525.
47. Vaughan DN, Rayner G, Tailby C, et al. MRI-negative temporal lobe epilepsy: a network disorder of neocortical connectivity. *Neurology*. 2016;87(18):1934-1942.
48. Doucet GE, Sharan A, Pustina D, et al. Early and late age of seizure onset have a differential impact on brain resting-state organization in temporal lobe epilepsy. *Brain Topogr*. 2015;28(1):113-126.
49. Yang Z, Choupan J, Reutens D, et al. Lateralization of temporal lobe epilepsy based on resting-state functional magnetic resonance imaging and machine learning. *Front Neurol*. 2015;6:184.
50. Larivière S, Rodríguez-Cruces R, Royer J, et al. Network-based atrophy modeling in the common epilepsies: A worldwide ENIGMA study. *Sci Adv*. 2020;6(47):eabc6457.

Figure and Table Captions

Figure 1: Top left shows a single frame of BOLD fMRI from a healthy control participant. Top right shows the same frame after noise injection via randomly permuting the BOLD time series within the highlighted circle. Bottom panel shows the BOLD time series from a single voxel (Raw) versus the same time series after noise injection or random permutation.

Figure 2. Grad-CAM for left and right seizure onset zones (radiological convention, slices shown are focused on the temporal lobes). Larger values represent stronger predictors. Depending on the lateralization, the 3DCNN learns features in the ipsilateral hippocampus, amygdala, insula, neocortical temporal lobe, orbitofrontal cortex and ipsilateral and contralateral parietal regions.

Figure 3: Average Grad-CAM weights and variance for top 12 structural Freesurfer regions. Left image shows the average Grad-CAM weights calculated for left TLE, and right image shows average Grad-CAM weights for right TLE. Anatomical regions most useful for classification were found in temporal, frontal, parietal, and insular regions.

Figure 4: Average Grad-CAM weights and variance for resting state networks. The RSNs most predictive overall include the default mode (DMN), dorsal attention (DAN), and medial temporal (MTL) networks. Left image shows the primary RSNs predictive for left TLE, including DMN, MTL, and frontal parietal network (FPN). Right image shows the primary RSN predictive for right TLE, which include DAN, DMN, and MTL. Note difference in scale of the two plots.

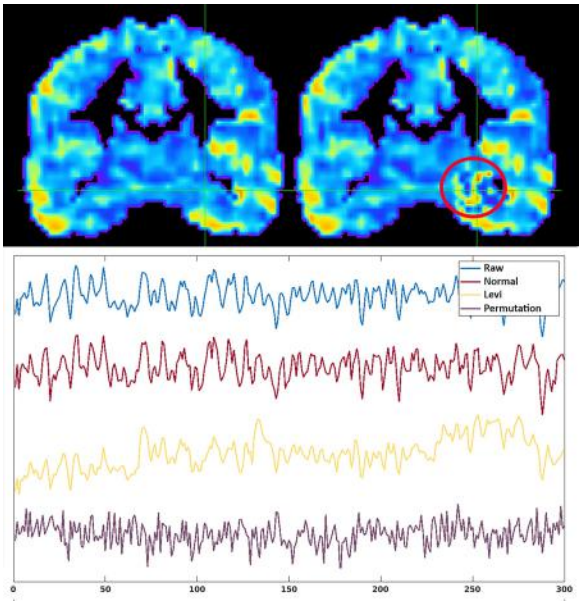
Figure 5. Mean Grad-CAM results (radiological convention) comparing Engel class 1 and Engel class >1 in left TLE patients. The mean Grad-CAM map for Engel class >1 shows more localized regions on the left contributing to the lateralization score with higher magnitude when compared to Engel class 1.

Table 1. Demographics

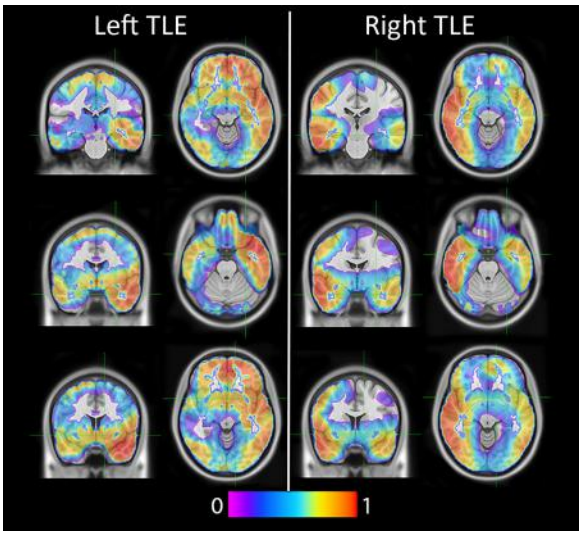
Table 1. Demographics

	Controls	Left TLE	Right TLE
N	2132	20	12
Age (mean±STD)	37.8±21.5	42.4±13.9	42.4±13.3
% Female	58%	55%	75%
Age at onset (mean±STD)	-	27.3 (17.5)	13.7 (14.9)
% MTS/Atrophy	-	40%	67%
% Surgical outcomes	-	45%	50%
% Engel Class 1	-	55%	100%

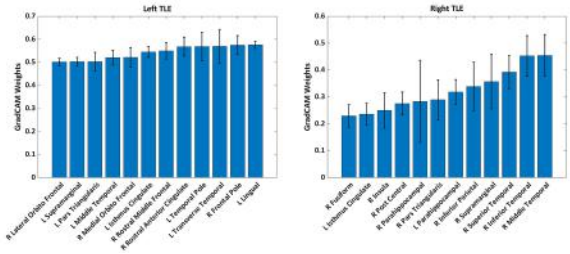
MTS, Mesial temporal sclerosis, STD, standard deviation



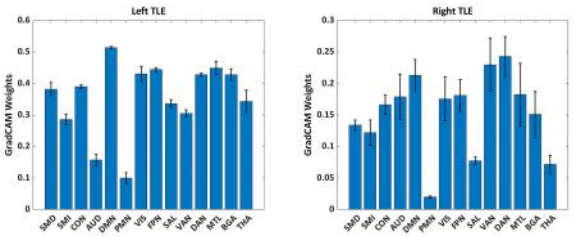
epi_17233_f1.tif



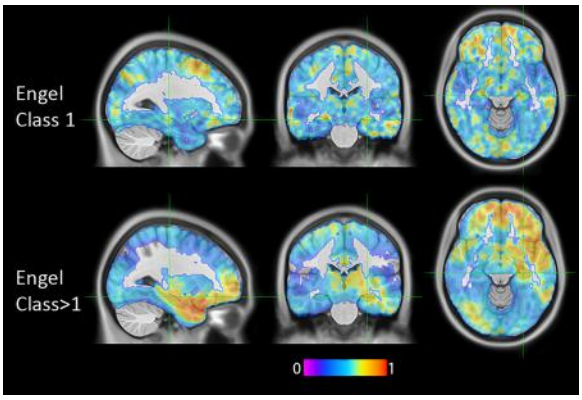
epi_17233_f2.tif



epi_17233_f3.tif



epi_17233_f4.tif



epi_17233_f5.tif