MRI-Based Machine Learning Prediction Framework to Lateralize Hippocampal Sclerosis in Patients With Temporal Lobe Epilepsy

Benoit Caldirou, PhD,* Niels A. Foit, MD,* Carlotta Mutti, MD, Fatemeh Fadaie, PhD, Ravnoor Gill, PhD, Hyo Min Lee, PhD, Theo Demerath, MD, Horst Urbach, MD, Andreas Schulze-Bonhage, MD, Andrea Bernasconi, MD, and Neda Bernasconi, MD, PhD

Correspondence

Dr. Bernasconi
neda@bic.mni.mcgill.ca

Neurology® 2021;97:e1583-e1593. doi:10.1212/WNL.00000000000012699

Abstract

Background and Objectives

MRI fails to reveal hippocampal pathology in 30% to 50% of temporal lobe epilepsy (TLE) surgical candidates. To address this clinical challenge, we developed an automated MRI-based classifier that lateralizes the side of covert hippocampal pathology in TLE.

Methods

We trained a surface-based linear discriminant classifier that uses T1-weighted (morphology) and T2-weighted and fluid-attenuated inversion recovery (FLAIR)/T1 (intensity) features. The classifier was trained on 60 patients with TLE (mean age 35.6 years, 58% female) with histologically verified hippocampal sclerosis (HS). Images were deemed to be MRI negative in 42% of cases on the basis of neuroradiologic reading (40% based on hippocampal volumetry). The predictive model automatically labeled patients as having left or right TLE. Lateralization accuracy was compared to electroclinical data, including side of surgery. Accuracy of the classifier was further assessed in 2 independent TLE cohorts with similar demographics and electroclinical characteristics (n = 57, 58% MRI negative).

Results

The overall lateralization accuracy was 93% (95% confidence interval 92%–94%), regardless of HS visibility. In MRI-negative TLE, the combination of T2 and FLAIR/T1 intensities provided the highest accuracy in both the training (84%, area under the curve [AUC] 0.95 ± 0.02) and validation (cohort 1 90%, AUC 0.99; cohort 2 76%, AUC 0.94) cohorts.

Discussion

This prediction model for TLE lateralization operates on readily available conventional MRI contrasts and offers gain in accuracy over visual radiologic assessment. The combined contribution of decreased T1- and increased T2-weighted intensities makes the synthetic FLAIR/T1 contrast particularly effective in MRI-negative HS, setting the basis for broad clinical translation.

Classification of Evidence

This study provides Class II evidence that in people with TLE and MRI-negative HS, an automated MRI-based classifier accurately determines the side of pathology.

MORE ONLINE

Class of Evidence

Criteria for rating therapeutic and diagnostic studies

[NPublic.org/coe](https://npublic.org/coe)

*These authors contributed equally to this work.

From the Neuroimaging of Epilepsy Laboratory (B.C., N.A.F., C.M., F.F., R.G., H.M.L., A.B., N.B.), McConnell Brain Imaging Center, Montreal Neurological Institute, McGill University, Quebec, Canada; and Departments of Neurosurgery (N.A.F.) and Neuroradiology (T.D., H.U.), Freiburg Medical Center, and Department of Neurology (A.S.-B.), University of Freiburg, Germany.

Go to [Neurology.org/N](https://neurology.org/N) for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

Glossary

AUC = area under the curve; **FLAIR** = fluid-attenuated inversion recovery; **FOV** = field of view; **HS** = hippocampal sclerosis; **LDA** = linear discriminant analysis; **LTLE** = left TLE; **ROC** = receiver operating characteristics; **ROI** = region of interest; **RTLE** = right TLE; **SEEG** = stereoencephalography; **TE** = echo time; **3D** = 3-dimensional; **TI** = inversion time; **TLE** = temporal lobe epilepsy; **TR** = repetition time; **2D** = 2-dimensional.

Many patients with medically refractory temporal lobe epilepsy (TLE) present with hippocampal sclerosis (HS).¹ MRI is instrumental for the identification of this pathology that may form the substrate of the epileptogenic focus, thus streamlining the presurgical evaluation.² The main imaging characteristics of HS are loss of hippocampal volume, often associated with hypointense T1- and hyperintense T2-weighted signal. The biological validity of these features has been established through combined MRI-histopathologic analyses showing that decreased cell density and gliosis correlate positively with atrophy³ and T2 hyperintensity,⁴ respectively. In clinical practice, however, MRI still fails to reveal hippocampal pathology in 30% to 50% of surgical candidates with unambiguous electroclinical evidence of TLE.⁵ This wide range may be at least partly attributable to suboptimal imaging protocols and limited specialized experience.⁵ In addition, while quantitative analyses, including hippocampal volumetry,⁶ voxel-based morphometry,⁷ T2 relaxometry,⁸ and measurements of fluid-attenuated inversion recovery (FLAIR) signal intensity,^{9,10} have been shown to be more sensitive compared to visual evaluation, they remain underused (see elsewhere¹¹ for review). The *in vivo* signature of HS is modulated by the severity of loss of neurons and gliosis,^{12,13} with subtle forms typified by isolated gliosis¹¹ often evading detection.^{1,14} Many patients with unrevealing MRI may thus undergo intracranial EEG, a procedure that carries risks similar to those of resective surgery¹⁵ and incurs high costs.¹⁶

Despite the large body of neuroimaging literature assessing hippocampal structural integrity in TLE, the vast majority of studies have addressed group-level changes.^{5,6,17} Individual analyses, on the other hand, have commonly exploited single contrasts normalized to the distribution of healthy controls and rarely addressed the challenge of lateralizing MRI-negative patients.^{6,18,19} Our purpose was to implement a machine learning framework to lateralize HS in individual patients relying on readily available conventional T1- and T2-weighted contrasts.²⁰ Because HS is typically characterized by T1-weighted hypointensity and T2-weighted hyperintensity, we also generated a synthetic contrast by dividing the FLAIR intensity by T1-weighted intensity, thereby maximizing their combined contributions to detect the full HS spectrum. We applied the classifier to MRI features of HS in TLE and assessed generalizability in 2 independent cohorts.

Methods

We present an algorithm for automated lateralization of the epileptogenic lesion in TLE, with the prediction model

labeling patients as have left TLE (LTLE) or right TLE (RTLE). In short, we trained a surface-based linear discriminant classifier using volume and signal intensity based on T1- and T2-weighted MRI and FLAIR/T1 of patients with histologically verified HS. Lateralization accuracy was compared to electroclinical data, including side of surgery. To address generalizability, accuracy was tested in 2 independent validation TLE cohorts with similar electroclinical and imaging characteristics. Classifiers were cross-validated with a 5-fold scheme with 100 repetitions. Evaluation performance was further assessed via receiver operating characteristics (ROC) curves and area under the curve (AUC).

Participants

Training Cohort

Sixty consecutive patients with TLE and validated HS were collected from 2010 to 2014 retrospectively. All had research-dedicated 3T MRI comprising high-resolution 3-dimensional (3D) T1-weighted, 3D FLAIR, and 2-dimensional (2D) T2-weighted images, as well as equivalent clinical MRI sequences as part of the presurgical evaluation. Patients had been investigated for drug-resistant epilepsy with a standard presurgical workup that include a neurologic examination, history of seizures, and EEG telemetry. Clinical neuroradiologic diagnosis of HS was based on morphologic anomalies of the hippocampus (atrophy, loss of internal structure, decreased T1, and increased T2). Side-by-side comparison of morphology, including shape, and signal was done on coronal images, while sagittal cuts yielded anteroposterior evaluation, thereby easing the visibility of distribution of signal anomalies.

Validation Cohorts

Validation cohorts comprised 56 patients with drug-resistant TLE collected from 2015 to 2017. The internal cohort comprised 43 consecutive cases seen at the Montreal Neurologic Hospital retrospectively (26 with histologically validated HS). The external cohort consisted of 14 patients with TLE from the Freiburg Epilepsy Center in Germany (13 with histologically validated HS). Patients in both cohorts underwent the same clinical and MRI evaluation as the training cohort. Imaging was done either on a 3T Siemens Trio or Prisma scanners (Siemens, Munich, Germany).

Standard Protocol Approvals, Registrations, and Patient Consents

The Research Ethics boards of the Montreal Neurologic Hospital and the Freiburg Medical Center gave approval of the study, and all participants gave informed consent.

MRI Imaging and Processing

In the training cohort and controls, images were acquired on a 3T TimTrio using a 32-arrays coil with the HARNESS-MRI protocol,²⁰ including 3D T1-weighted magnetization-prepared rapid acquisition with gradient echo (repetition time [TR] 2,300 milliseconds, echo time [TE] 2.98 milliseconds, inversion time [TI] 900 milliseconds, flip angle 9°, matrix 256×256 , field of view [FOV] 256×256 mm, yielding 1-mm^3 voxels; 6.35 minutes), 3D FLAIR (TR 5,000 milliseconds, TE 390 milliseconds, TI 1,800 milliseconds, matrix 230×230 , FOV 207×207 mm, 0.9 mm^3 ; 6.22 minutes), and 2D turbo spin echo T2 sequence (TR 10,810 milliseconds, TE 81 milliseconds, flip angle 119°, matrix 512×512 , FOV 203×203 , $0.4 \times 0.4 \times 2.0\text{ mm}^3$; 5.5 minutes). The acquisition parameters were similar for the validation cohorts.

In all participants, T1-weighted, FLAIR, and T2-weighted images underwent field nonuniformity correction, followed by a standardization of signal, as well as alignment to the International Consortium of Brain Mapping-152 template.²¹ FLAIR and T2-weighted images were registered to T1 MRI in Montreal Neurological Institute space. A surface-based multitemplate algorithm²² automatically segmented the hippocampus in the CA1 to CA3, CA4 to DG, and subiculum²³; this algorithm has shown excellent Dice overlap indices with manual segmentations.²²

Classifier Design

Data Sampling

We mapped the medial sheet, namely a surface along the central axis of each hippocampal subfield; this allowed extracting features minimizing effects of partial volume.²⁴ We obtained the following features on the medial sheets:

1. Columnar volume. Neuronal loss is associated with atrophy on MRI, which we estimated by calculating columnar volume as done previously.²⁴
2. T2 signal intensity. Gliosis is characterized by increased T2 signal intensity. We computed the relative intensity of T2 at every voxel.²⁵
3. FLAIR/T1 intensity. To maximize the detection of HS, typically characterized by T1-hypointensity and T2-weighted hyperintensity, we generated a synthetic contrast by dividing the FLAIR by T1-weighted intensity, thereby maximizing their combined contribution.

After z normalizing each feature with respect to healthy controls, we generated asymmetry maps computed as $[2 \times (\text{left} - \text{right}) / (\text{left} + \text{right})]$, which served as inputs to the algorithm. The gold standard for training and cross-validation consisted of patients with TLE with histologically validated HS (training cohort).

Architecture

Our algorithm aims at lateralizing the epileptogenic lesion by assigning patients to either LTLE or RTLE (Figure 1A).

Lateralization is formulated as a classification task, leveraging linear discriminant analysis (LDA) as the classifier. LDA requires virtually no parameter tuning while operating efficiently in low-dimensional space with a limited number of features, as in our case, obviating the need for more complex or nonlinear algorithms with a significant time requirement. The training procedure consists of 2 components. First, we generate the optimal region of interest (ROI), namely a spatial constraint for feature sampling and averaging derived from the asymmetry maps of columnar volume and T2 and FLAIR/T1 intensities. More specifically, paired t tests (or Hotelling T^2 for feature combinations) contrasted corresponding vertices of the ipsilateral and contralateral hippocampal subfields (relative to the epileptogenic focus) to highlight regions exhibiting the largest feature asymmetry. The resulting t map was then thresholded from zero to the highest absolute t statistic to generate binarized t maps. The choice of the threshold determined the spatial extent of the binarized t map, and consequently the discriminative ability of the averaged asymmetry features. Therefore, this binarized t map (or ROI) served as a hyperparameter optimized through a nested cross-validation procedure. This process selected the most performant model²⁶ (i.e., the optimal ROI or threshold; Figure 1B) while mitigating the propagation of ground truth information across folds (also known as data leakage²⁷). Subsequently, for each feature, values sampled on the asymmetry maps (constrained by the optimal ROI) were averaged across the hippocampus. Last, these values served as inputs to the LDA, which determined laterality (i.e., LTLE or RTLE) on the basis of the learned statistical patterns.

Predictors Considered for Modeling

To evaluate the differential impact of features and their combinations, we tested the lateralization performance of the classifier when using (1) columnar volume; (2) T2-weighted intensity; (3) FLAIR/T1 intensity; and (4) a combination of T2 and FLAIR/T1. Notably, volumetry was excluded from the combinatorial analysis because it is not expected to be discriminative for MRI-negative patients because their hippocampal size is generally within the normative range of healthy controls.²⁵

Performance Evaluation

For the training cohort, performance was assessed through a 5-fold (nested) cross-validation repeated 100 times. Stratification ensured that each fold had proportional representation of both LTLE and RTLE. Briefly, after the training cohort was randomly split into 5 folds (or partitions), the classifier was trained on 4 and tested iteratively on the one held out until all folds had served as a test set; this procedure was repeated 100 times. To assess generalizability, algorithmic performance was tested on 2 validation cohorts. To guarantee the highest confidence, we trained the classifier on a random sample comprising 80% of patients from the training cohort, repeating the process 100 times.

In designing this classifier, our objective was to determine the side of the pathology (not whether it is present or absent). Our primary performance validation metric was

Figure 1 Classifier Design

(A) Training. Objective of training was to define an optimal region of interest (ROI) used to sample MRI features of hippocampal sclerosis. (1) For each feature in the training set, paired t tests compared corresponding vertices of the ipsilateral and contralateral hippocampal subfields, z scored with respect to healthy controls (only maps of T2-weighted anomalies shown). (2) Resulting group-level t map was exhaustively thresholded from 0 to the highest value and binarized. (3) For each threshold, we overlaid the binarized t map on the asymmetry map of each individual and computed the average across subfields. (4) We then trained 1 linear discriminant analysis (LDA) classifier per threshold and retained the model yielding the highest lateralization accuracy (here, LDA model 3, surrounded by the black dotted box) and used it to test the classifier. (B) Statistical parametric anatomical map of optimal ROIs. For each modality, maps show the vertex-wise group-level probability of anomalies (optimal ROI) over 100 repetitions of the 5-fold cross-validation determined during training. FLAIR = fluid-attenuated inversion recovery; LTLE = left TLE; RTLE = right TLE; TLE = temporal lobe epilepsy.

thus accuracy, which reflects the average of 2 classes (RTLE and LTLE); we did not intend to evaluate the class dichotomy of a patient with TLE vs a healthy control. We also obtained ROC and AUC curves. Lateralization accuracy and AUC were measured for each 5-fold validation repetition and averaged across them. The Brier score was used as a measure of calibration²⁸; values close to zero signify a well-calibrated classifier. Comparisons among experiments were assessed with the Friedman test with Bonferroni correction.

Statistical Analysis

Group Analysis

Statistical analysis was carried out with SurfStat (Matlab, MathWorks, Natick, MA). For each participant, we first z scored vertex-wise values (columnar volume, normalized T2 intensity, and FLAIR/T1) based on healthy controls. We then sorted the left and right values into ipsilateral and contralateral with respect to the focus. The Student test assessed differences between patients and controls, correcting at a family-wise error of $p_{\text{FWE}} < 0.05$ using random field theory.

Data Availability

The source codes for (1) generating blade surfaces (as described in the Data Sampling section), (2) data sampling (intersection of blade surfaces and volumes), (3) computing columnar volumes, and (4) training and testing the classifier are available from the corresponding author. We also make available the optimal ROI and the pretrained model (based on the training cohort) to enable lateralization prediction on a test participant without the need to collect data to train the classifier.

Results

Table 1 details clinical and demographic features of the training and validation cohorts.

Clinical, Demographics, and Imaging Characteristics

Training Cohort

In 42 patients, the focus side was determined by EEG monitoring with scalp electrodes showing unequivocal temporal

Table 1 Training and Validation Cohorts Characteristics

	Age, y	Female, %	SEEG, %	TLE (L/R), n	HS/G, n	Engel class I, %
Training						
MRI positive (n = 36)	36 ± 11	53	17	17/19	36/0	89
MRI negative (n = 24)	35 ± 11	67	50 ^a	12/12	5/19 ^a	67
Validation 1						
MRI positive (n = 16)	39 ± 12	75	19	10/6 ^a	13/0	92
MRI negative (n = 27)	36 ± 10	60	52 ^a	20/7 ^a	3/10 ^a	70
Validation 2						
MRI positive (n = 8)	33 ± 13	50	25	4/4	8/0	100
MRI negative (n = 6)	42 ± 11	67	67	1/5	0/5 ^a	100

Abbreviations: HS/G = hippocampal sclerosis/isolated gliosis; SEEG = stereoencephalography; TLE = temporal lobe epilepsy. L/R is EEG lateralization.

^a Higher proportions in a given category.

lobe seizures onset (and >70% of spikes); in cases with nonlocalized seizure onset or rapid interhemispheric seizure spread (n = 18), lateralization was established with stereoencephalography (SEEG). Accordingly, patients were dichotomized into LTLE (n = 29, 17 female patients, mean ± SD age 35.6 ± 11 years, range 18–59 years) and RTLE (n = 31, 18 female patients, age 35.5 ± 11 years, range 17–62 years). As per the reading of the neuroradiologist, 35 patients (35 of 60 = 58%) had ipsilateral atrophy of the hippocampus together with T2 hypersignal (MRI positive), while the MRI was reported as unremarkable in 25 (42%, MRI negative). No other anomalies were seen. Performing volumetry had minimal impact, with only a single MRI-negative patient becoming MRI-positive with an ipsilateral hippocampal volume reduction of –2.5 SD below the mean of healthy controls, bringing the total count to 36 (60%) MRI-positive and 24 (40%) MRI-negative patients.

According to a histopathologic review of the resected specimen, 40 patients had severe neuronal cell loss and astrogliosis: 23 in CA1 to CA3 and CA4 (International League Against Epilepsy HS-1), 9 CA1 predominant (HS-2), and 9 CA4 (HS-3); 19 patients showed isolated gliosis.¹ Notably, all MRI-positive patients had HS, while MRI-negative patients presented with both mild HS (5 of 24 = 21%; 3 type 2 and 2 type 3) and isolated gliosis (19 of 24 = 79%). At a follow-up of >2 years, Engel class I was reported in 48 (80%) patients, Engel class II in 9 (15%), Engel class III in 3 (3.3%), and Engel class IV in 1 (1.7%). Thirty-six healthy individuals (18 female patients, age 32.2 ± 7.3 years, range 23–53 years) formed the control group.

Validation Cohorts

On the basis of the same criteria as in the training cohort, patients were dichotomized into LTLE (n = 35, 25 female patients, mean ± SD age 37.2 ± 11 years, range 19–58 years)

and RTLE (n = 22, 11 female patients, age 36.9 ± 12 years, range 18–54 years) based on scalp EEG (n = 36) and SEEG (n = 21). Twenty-four patients (24 of 57 = 42%) has ipsilateral hippocampal volume reduction and high T2 (MRI positive), while the MRI was reported as unremarkable in 33 (58%, MRI negative). No patient had hippocampal atrophy on volumetry. Thirty-nine patients had surgery. Histopathology showed severe HS in 24 (HS-1 in 15, HS-2 in 6, HS-3 in 3) and isolated gliosis in 15.¹

Comparisons

We observed no difference among cohorts for age (1-way analysis of variance, $p = 0.73$), sex (χ^2 test, $\chi^2 = 0.57$, $p = 0.75$), proportion of MRI-positive and MRI-negative patients ($\chi^2 = 5.43$, $p = 0.07$), and surgical outcome ($\chi^2 = 3.59$, $p = 0.17$). In all 3 cohorts, the proportion of isolated gliosis was higher than HS in MRI-negative patients (training cohort: $\chi^2 = 38.13$, $p < 0.001$; validation cohort 1: $\chi^2 = 13.16$, $p < 0.001$; validation cohort 2: $\chi^2 = 9.12$, $p = 0.002$). The proportion of patients who underwent SEEG was higher among MRI-negative patients with respect to MRI-positive patients in the training cohort ($\chi^2 = 7.62$, $p = 0.006$) and validation cohort 1 ($\chi^2 = 4.61$, $p = 0.032$). The proportion with LTLE was higher than the proportion with RTLE in validation cohort 1 ($\chi^2 = 6.96$, $p = 0.03$).

Group Analysis

Compared to the healthy controls (Figure 2), patients exhibited diffuse ipsilateral atrophy across all subfields (CA1–CA3: $t = -2.5$, $p_{FWE} < 0.0001$; subiculum: $t = -2.2$, $p_{FWE} < 0.0001$; CA4–DG: $t = 2.3$, $p_{FWE} < 0.0001$). Moreover, marked ipsilateral T2 hypersignal signal was present in CA1 to CA3 ($t = 2.8$, $p_{FWE} < 0.0001$) and CA4 to DG ($t = 3.9$, $p_{FWE} < 0.0001$). FLAIR/T1 was also increased across all subregions (CA1–CA3: $t = 3.5$, $p_{FWE} < 0.0001$; CA4–DG: $t = 4.3$, $p_{FWE} = 0.004$; subiculum: $t = 2.4$, $p_{FWE} < 0.0001$), with additional

subtle increases contralaterally. Anomalies in MRI-positive patients had similar distributions across subfields, albeit more severe (volume: $t = -3.9$ to -4.4 , $p_{\text{FWE}} < 0.0001$; T2 signal: $t = 3.6$ – 5.4 , $p_{\text{FWE}} < 0.0001$; FLAIR/T1: $t = 2.7$ – 6.0 , $p_{\text{FWE}} < 0.002$). Conversely, MRI-negative patients did not show any volumetric alteration but subtle T2 increases in the ipsilateral CA4 to DG ($t = 2.5$, $p_{\text{FWE}} < 0.0001$) and CA1 through CA3 ($t = 2.1$, $p_{\text{FWE}} = 0.014$), as well as FLAIR/T1 hyperintensities along all subfields (CA1 through CA3: $t = 2.5$, $p_{\text{FWE}} < 0.0001$; CA4–DG: $t = 2.9$, $p_{\text{FWE}} = 0.005$; subiculum: $t = 2.1$, $p_{\text{FWE}} < 0.0001$).

Performance Evaluation

In the training cohort, the overall lateralization accuracy based on individual features was similar for FLAIR/T1 ($85 \pm 3\%$) and T2 signal ($86 \pm 2\%$), which were superior to that of volumetry ($77 \pm 3\%$; $p < 0.05$). The combination of T2 and FLAIR/T1 yielded the best overall performance of $93 \pm 2\%$, with an accuracy of $84 \pm 5\%$ in MRI-negative and $100 \pm 1\%$ in MRI-positive TLE (Table 2). Notably, the lateralization of the classifier was consistently correct across 90 of 100 iterations in both MRI-negative and MRI-positive patients. AUC metrics (Table 3) confirmed the discriminative capability of FLAIR/T1 (0.80 ± 0.05) and T2 signal (0.79 ± 0.03) over volumetry (0.46 ± 0.06 , $p_{\text{Bonferroni}} < 0.05$) in MRI-negative patients; the combination of T2 and FLAIR/T1 yielded the highest AUC (0.95 ± 0.03 , $p_{\text{Bonferroni}} < 0.05$). The Brier score demonstrated better calibration of the classifier when the combination of T2 and FLAIR/T1 was used relative to other scenarios (0.097 ± 0.016 , $p_{\text{Bonferroni}} < 0.05$). In the validation cohorts, the combination of T2 and FLAIR/T1 also yielded the best overall performance with $>90\%$ lateralization accuracy. Moreover, AUC showed the discriminative capability of this combination in MRI-negative patients (1.00 ± 0.00 and 0.94 ± 0.12 for

validation cohorts 1 and 2, respectively). The Brier score reflected the improved calibration when combining T2 and FLAIR/T1 for validation cohort 1 relative to single features (0.069 ± 0.015 , $p_{\text{Bonferroni}} < 0.05$), while in validation cohort 2, similar scores were obtained for combined T2 and FLAIR/T1 (0.152 ± 0.053) and FLAIR/T1 (0.160 ± 0.020). Examples of lateralization predictions are shown in Figure 3. Figure 4 highlights the robustness of combining T2 and FLAIR/T1 to variability in the training dataset, as evidenced by the low variance of the average ROC curve.

Discussion

In epilepsy surgery, converging independent diagnostic tests aim at the lateralization of the epileptic focus based on neurophysiology and MRI localization of the epileptogenic lesion. When either of these 2 diagnostic pillars is unclear or unrevealing, additional investigations are needed. Invasive EEG recordings may clarify the focus lateralization, for example, in cases with nonlocalized seizure onset or rapid interhemispheric seizure spread. Moreover, MRI postprocessing and machine learning may clarify lesion location in cases with normal MRI. Here, we present a decision-support system to lateralize HS in TLE using structural MRI. The classifier relies on automatically segmented hippocampal subfields and feature sampling derived from conventional T1- and T2-weighted contrasts available on most magnetic resonance scanners.²⁰ Our method yields an overall accuracy of $>90\%$ regardless of the degree of HS visibility on conventional MRI. It is important to note that it lateralizes MRI-negative TLE with $>80\%$ accuracy, offering considerable gain over visual radiologic assessment.

Figure 2 Group-Level Findings

Differences in columnar volume, T2 intensity, and fluid-attenuated inversion recovery (FLAIR)/T1 intensity between patients and controls are mapped on the hippocampal subfield surfaces. Results are corrected with random field theory. Red and blue indicate increases and decreases, respectively (scaled by Cohen's d effect size).

Table 2 Lateralization Performance (Accuracy) Across Training and Validation Cohorts

	Volume, %	T2 intensity, %	FLAIR/T1, %	T2 + FLAIR/T1, %
Training				
MRI positive (n = 36)	97 ± 2 (97–98)	96 ± 2 (96–96)	95 ± 2 (94–95)	100 ± 1 ^a (100–100)
MRI negative (n = 24)	45 ± 7 (44–47)	71 ± 5 ^a (70–72)	71 ± 5 ^a (70–73)	84 ± 5 ^a (83–85)
Validation 1				
MRI positive (n = 16)	95 ± 3 (95–96)	96 ± 3 (95–96)	100 ± 2 ^a (99–100)	100 ± 0 ^a (100–100)
MRI negative (n = 27)	61 ± 8 (59–62)	78 ± 4 (78–79)	88 ± 4 ^a (88–89)	90 ± 4 ^a (90–91)
Validation 2				
MRI positive (n = 8)	99 ± 5 (98–100)	100 ± 0 (100–100)	99 ± 3 (99–100)	100 ± 0 (100–100)
MRI negative (n = 6)	55 ± 11 (53–57)	49 ± 5 (48–50)	70 ± 7 ^a (69–71)	76 ± 13 ^a (73–78)

(n) = sample size.

The 95% confidence intervals are shown in parentheses.

^a Increased/best lateralization accuracy with respect to at least 1/any model (Friedman $p < 0.05$ after correction).

Our purpose was to develop a prediction model with the potential to be implemented into clinical practice. To ensure a valid biological foundation, we trained the LDA algorithm on well-known, histologically validated features of HS. When using a conventional z-score comparison with healthy controls, we were able to lateralize pathology in only a fraction (<15%) of patients with MRI-negative TLE using volumetry, T2, or FLAIR/T1. This disparity in performance relative to machine learning can be explained mainly by the use of a spatial ROI, allowing sampling of features relevant to lateralization. This optimal ROI relies on the population-wise feature difference between the ipsilateral and contralateral hippocampal subfields, targeting regions with the most marked interhemispheric asymmetries. Through the training procedure, the extent of the ROI was adjusted iteratively,

thereby boosting lateralization accuracy in single cases. In contrast, a z-score normalization of the features of interest across the whole hippocampus (or each subfield) would likely include nonpertinent information, leading to suboptimal lateralization accuracy. More recent work based on non-conventional contrasts, including diffusion²⁹ and network parameters,^{30–32} has targeted whole-brain anomalies in patients with MRI-positive TLE. Only 2 previous studies have addressed the lateralization challenge in both MRI-positive and MRI-negative TLE,^{33,34} operating on T1-derived volumetry with groups predefined by side and visibility of hippocampal atrophy on MRI. In 1 study,³³ besides the lack of histopathologic confirmation, anatomic structures identifying TLE groups were mainly outside the mesiotemporal lobe, different across groups, and difficult to interpret, particularly

Table 3 Lateralization Performance (AUC) Across Cohorts

	Volume	T2 intensity	FLAIR/T1	T2 + FLAIR/T1
Training				
MRI positive (n = 36)	1.00 ± 0.01 (1.00–1.00)	1.00 ± 0.00 (1.00–1.00)	0.98 ± 0.01 (0.98–0.99)	1.00 ± 0.00 (1.00–1.00)
MRI negative (n = 24)	0.46 ± 0.06 (0.45–0.47)	0.79 ± 0.03 ^a (0.79–0.80)	0.80 ± 0.05 ^a (0.78–0.81)	0.95 ± 0.03 ^a (0.95–0.96)
Validation 1				
MRI positive (n = 16)	1.00 ± 0.00 (1.00–1.00)	1.00 ± 0.00 (1.00–1.00)	1.00 ± 0.00 (1.00–1.00)	1.00 ± 0.00 (1.00–1.00)
MRI negative (n = 27)	0.65 ± 0.06 (0.64–0.66)	0.91 ± 0.02 ^a (0.91–0.92)	0.93 ± 0.04 ^a (0.93–0.94)	1.00 ± 0.00 ^a (1.00–1.00)
Validation 2				
MRI positive (n = 8)	1.00 ± 0.03 (0.98–1.00)	1.00 ± 0.00 (1.00–1.00)	1.00 ± 0.00 (1.00–1.00)	1.00 ± 0.00 (1.00–1.00)
MRI negative (n = 6)	0.61 ± 0.18 (0.57–0.65)	0.65 ± 0.16 (0.62–0.68)	0.76 ± 0.10 ^a (0.74–0.78)	0.94 ± 0.12 ^a (0.91–0.96)

Abbreviations: AUC = area under the curve; FLAIR = fluid-attenuated inversion recovery.

The 95% confidence intervals are shown in parentheses.

^a Increased/best lateralization accuracy with respect to at least 1/any model (Friedman $p < 0.05$ corrected for multiple comparisons).

Figure 3 Individual Lateralization Prediction

Examples of lateralization prediction in 2 patients with MRI-positive right temporal lobe epilepsy (TLE) (A) and MRI-negative left TLE (B). For each case, coronal sections of the T1-weighted and T2-weighted MRI and the synthetic fluid-attenuated inversion recovery (FLAIR)/T1 contrast (right is right on images) are shown together with the automatically generated asymmetry maps for columnar volume, T2-weighted, and FLAIR/T1 intensities. On each map, dotted line corresponds to the coronal MRI section, and optimal ROI (Figure 1) is outlined in black.

in MRI-negative patients. While methodologically appropriate, such a design may be unsatisfactory in a real-world scenario of previously unseen cases.

Our study presents several strengths. First, it uses conventional contrasts.²⁰ Second, we targeted morphologic and signal alterations of the hippocampus, the surgically amenable substrate of mesial TLE. Third, effects of partial volume were minimized by surface-based image processing. For individual prediction, we opted for an LDA, a robust and easy-to-interpret classifier.³⁵ The classifier operated in regions that were identical in both MRI-negative and MRI-positive TLE. From a statistical standpoint, to mitigate optimistic estimates of lateralization accuracy (otherwise known as model overfitting) that would occur with a conventional leave-one-out scheme, we kept the training and testing datasets separate using a cross-validation with 5 folds and 100 repetitions. The effectiveness and generalizability of our algorithm are also evidenced by the high degree of consistency across repetitions, which is further supported by high AUC values in both the training and validation cohorts. Indeed, when the classifier was applied to independent datasets of unseen cases including both MRI-positive and MRI-negative patients, accuracy reached 90% correct lateralization. Despite excellent performance, our method presents some limitations. Notably, contrary to previous work,³⁶ our design does not allow discriminating patients from controls. Moreover, our sample size for training may be considered relatively small with respect to general machine learning standards. Nevertheless, our

numbers compare favorably to work in epilepsy. Indeed, prior studies analyzing imaging acquired at 3T were based on a maximum of 80 individuals when both training and testing sets were combined.^{33,36} In contrast, our study targeted a significantly larger cohort of patients with 60 individuals for training and 57 for validation.

High performance and generalizability of the classifier across cohorts, scanners, and parameters set the basis for wide translation. Nevertheless, successful integration into clinical practice rests on key requirements. In line with recent educational initiatives,^{20,37} clinicians should develop competencies in neuroimaging spanning from basic visual diagnostics to the interpretation of advanced postprocessing methods and machine learning. The availability of the source code is intended to foster novel synergies between engineers and clinicians within the epilepsy community and should be regarded as a first step toward an online application or a tool integrated into patients' electronic health records.

The imaging correlates of HS associating neuronal loss and gliosis have long been established.³ At the group level, all features in our MRI-positive patients were significantly abnormal, including extensive ipsilateral hippocampal volume loss and increased T2-weighted and FLAIR/T1 signal intensities compared to healthy controls. Therefore, volumetry was as efficient to map pathology and to lateralize individual patients as any other intensity features. Conversely, MRI-negative cases, the vast majority with isolated gliosis, exhibited

Figure 4 ROC Curves in MRI-Negative TLE

Receiver operating characteristic (ROC) curves based on lateralization posterior probabilities from the trained linear discriminant analysis model are shown for the training and validation cohorts. The x- and y-axes represent the lateralization false positive rate (FPR) and true positive rate (TPR), respectively. Dotted blue lines represent individual curves drawn from each validation repetition; thick lines (green, columnar volume; purple, T2-weighted intensity; orange, fluid-attenuated inversion recovery [FLAIR]/T1 intensity; red, T2 + FLAIR/T1) show average ROC curves across iterations; dashed black lines correspond to a random classifier. Average area under the curve (AUC) is indicated. Right-most panels show the overlay of all curves for each cohort to facilitate comparisons.

only subtle ipsilateral signal anomalies and no volumetric alterations (aside from 1 patient with reduced volume). Expectedly, when applying machine learning to volumetry alone, performance was at a chance level, whereas a combination of intensity features derived from T2- and T1-weighted MRI (i.e., T2 together with FLAIR/T1) outperformed any unimodal contrast, thereby offering substantial gain over expert visual assessments. While our results are in agreement with previous studies demonstrating the value of T2 signal in the detection of gliosis,^{4,19,38} FLAIR/T1 was more effective in lateralizing the seizure focus. A possible biological explanation may be that in MRI-negative TLE severe gliosis coexists with subtle neuronal loss below the 10% sensitivity threshold of qualitative histopathology.³⁹ In addition, because 30% of these patients had mild HS, it is plausible that 3D FLAIR/T1 maximized the combined sensitivity of both contrasts to detect the full spectrum of hippocampal pathology. From a practical standpoint, our results also suggest that FLAIR/T1 may be a good alternative to 2D coronal T2-weighted images more prone to movement artifacts.

Nondiagnostic MRIs have led to an increase in invasive studies. By offering noninvasive decision support, advanced imaging analysis circumvents some of the limitations and risks related to invasive diagnostics,⁴⁰ possibly reducing the need for prolonged costly hospitalizations. Conversely,

because the presurgical evaluation of drug-resistant TLE is multidisciplinary, an MRI-derived binary lateralization outcome (right vs left) may not be sufficient per se for surgical decision-making. However, the increased availability of imaging-derived classification algorithms ought to pave the way for systems that integrate diverse sources of evidence, including other imaging modalities such as PET, as well as electroclinical data to increase diagnostic yield and certainty. Last, the methodology presented in this study may be expanded to other epilepsy syndromes associated with HS, including cortical developmental malformations.

Study Funding

This work was funded by the Canadian Institutes of Health Research (CIHR MOP-57840 to A.B. and CIHR MOP-123520 to N.B.), Natural Sciences and Research Council (Discovery-243141 to A.B. and 24779 to N.B.), Epilepsy Canada (Jay & Aiden Barker Breakthrough Grant in Clinical & Basic Sciences to A.B.), and Canada First Research Excellence Fund (HBHL-1a-5a-06 to N.B.). Salary supports were provided by the German Research Foundation (FO996/1-1) and the International League Against Epilepsy (N.A.F.), Savoy Foundation for Epilepsy (F.F., H.M.L.), Fonds de Recherche en Sante Quebec (R.G.), and Lloyd Carr-Harris Foundation (B.C.).

Disclosure

B. Caldairou, N. A. Foit, C. Mutti, F. Fadaie, R. Gill, H. M. Lee, T. Demerath, H. Urbach, A. Schulze-Bonhage, A. Bernasconi, and N. Bernasconi report no disclosure relevant to the manuscript. Go to [Neurology.org/N](https://www.neurology.org/N) for full disclosures.

Publication History

Received by *Neurology* December 1, 2020. Accepted in final form July 30, 2021.

Appendix Authors

Name	Location	Contribution
Benoit Caldairou, PhD	Neuroimaging of Epilepsy Laboratory, McConnell Brain Imaging Center, Montreal Neurological Institute, McGill University, Quebec, Canada	Drafting/revision of the manuscript for content, including medical writing for content; study concept or design; analysis or interpretation of data; statistical analysis
Niels A Foit, MD	Neuroimaging of Epilepsy Laboratory, McConnell Brain Imaging Center, Montreal Neurological Institute, McGill University, Quebec, Canada; Freiburg Medical Center, Department of Neurosurgery, University of Freiburg, Germany	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; analysis or interpretation of data
Carlotta Mutti, MD	Neuroimaging of Epilepsy Laboratory, McConnell Brain Imaging Center, Montreal Neurological Institute, McGill University, Quebec, Canada	Drafting/revision of the manuscript for content, including medical writing for content; Analysis or interpretation of data
Fatemeh Fadaie, PhD	Neuroimaging of Epilepsy Laboratory, McConnell Brain Imaging Center, Montreal Neurological Institute, McGill University, Quebec, Canada	Drafting/revision of the manuscript for content, including medical writing for content; analysis or interpretation of data
Ravnoor Gill, PhD	Neuroimaging of Epilepsy Laboratory, McConnell Brain Imaging Center, Montreal Neurological Institute, McGill University, Quebec, Canada	Drafting/revision of the manuscript for content, including medical writing for content; study concept or design; analysis or interpretation of data
Hyo Min Lee, PhD	Neuroimaging of Epilepsy Laboratory, McConnell Brain Imaging Center, Montreal Neurological Institute, McGill University, Quebec, Canada	Drafting/revision of the manuscript for content, including medical writing for content; analysis or interpretation of data
Theo Demerath, MD	Freiburg Medical Center, Department of Neuroradiology, University of Freiburg, Germany	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data
Horst Urbach, MD	Freiburg Medical Center, Department of Neuroradiology, University of Freiburg, Germany	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data
Andreas Schulze-Bonhage, MD	Department of Neurology, University of Freiburg, Germany	Drafting/revision of the manuscript for content, including medical writing for content; analysis or interpretation of data

Appendix (continued)

Name	Location	Contribution
Andrea Bernasconi, MD	Neuroimaging of Epilepsy Laboratory, McConnell Brain Imaging Center, Montreal Neurological Institute, McGill University, Quebec, Canada	Drafting/revision of the manuscript for content, including medical writing for content; study concept or design; analysis or interpretation of data
Neda Bernasconi, MD, PhD	Neuroimaging of Epilepsy Laboratory, McConnell Brain Imaging Center, Montreal Neurological Institute, McGill University, Quebec, Canada	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design; analysis or interpretation of data

References

- Blümcke I, Thom M, Aronica E, et al. International consensus classification of hippocampal sclerosis in temporal lobe epilepsy: a Task Force report from the ILAE Commission on Diagnostic Methods. *Epilepsia*. 2013;54(7):1315-1329.
- Jones AL, Cascino GD. Evidence on use of neuroimaging for surgical treatment of temporal lobe epilepsy: a systematic review. *JAMA Neurol*. 2016;73(4):464-470.
- Cascino GD, Jack CR, Parisi JE, et al. Magnetic resonance imaging-based volume studies in temporal lobe epilepsy: pathological correlations. *Ann Neurol*. 1991;30(1):31-36.
- Peixoto-Santos JE, Kandratavicius L, Velasco TR, et al. Individual hippocampal subfield assessment indicates that matrix macromolecules and gliosis are key elements for the increased T2 relaxation time seen in temporal lobe epilepsy. *Epilepsia*. 2017;58(1):149-159.
- Cendes F, Theodore WH, Brinkmann BH, Sulc V, Cascino GD. Neuroimaging of epilepsy. *Handb Clin Neurol*. 2016;136:985-1014.
- Vos SB, Winston GP, Goodkin O, et al. Hippocampal profiling: localized magnetic resonance imaging volumetry and T2 relaxometry for hippocampal sclerosis. *Epilepsia*. 2020;61(2):297-309.
- Riederer F, Seiger R, Lanzenberger R, et al. Voxel-based morphometry: from hype to hope: a study on hippocampal atrophy in mesial temporal lobe epilepsy. *AJNR Am J Neuroradiol*. 2020;41(6):987-993.
- Winston GP, Vos SB, Burdett JL, Cardoso MJ, Ourselin S, Duncan JS. Automated T2 relaxometry of the hippocampus for temporal lobe epilepsy. *Epilepsia*. 2017;58(9):1645-1652.
- Focke NK, Bonelli SB, Yogarajah M, Scott C, Symms MR, Duncan JS. Automated normalized FLAIR imaging in MRI-negative patients with refractory focal epilepsy. *Epilepsia*. 2009;50(6):1484-1490.
- Beheshti I, Sone D, Maikusa N, et al. FLAIR-wise machine-learning classification and lateralization of MRI-negative (18)F-FDG PET-positive temporal lobe epilepsy. *Front Neurol*. 2020;11:580713.
- Martin P, Bender B, Focke NK. Post-processing of structural MRI for individualized diagnostics. *Quant Imaging Med Surg*. 2015;5(2):188-203.
- Rudie JD, Colby JB, Salamon N. Machine learning classification of mesial temporal sclerosis in epilepsy patients. *Epilepsy Res*. 2015;117:63-69.
- Bernhardt BC, Hong S-J, Bernasconi A, Bernasconi N. Magnetic resonance imaging pattern learning in temporal lobe epilepsy: classification and prognostics. *Ann Neurol*. 2015;77(3):436-446.
- Hattingen E, Enkirch SJ, Jurcoane A, et al. Hippocampal "gliosis only" on MR imaging represents a distinct entity in epilepsy patients. *Neuroradiology*. 2018;60(2):161-168.
- Tanriverdi T, Ajlan A, Poulin N, Olivier A. Morbidity in epilepsy surgery: an experience based on 2449 epilepsy surgery procedures from a single institution. *J Neurosurg*. 2009;110(6):1111-1123.
- Muhlhofer W, Tan Y-L, Mueller SG, Knowlton R. MRI-negative temporal lobe epilepsy: what do we know? *Epilepsia*. 2017;58(5):727-742.
- Caciagli L, Bernasconi A, Wiebe S, Koepp MJ, Bernasconi N, Bernhardt BC. A meta-analysis on progressive atrophy in intractable temporal lobe epilepsy: time is brain? *Neurology*. 2017;89(5):506-516.
- Mueller SG, Laxer KD, Barakos J, Cheong I, Garcia P, Weiner MW. Subfield atrophy pattern in temporal lobe epilepsy with and without mesial sclerosis detected by high-resolution MRI at 4 Tesla: preliminary results. *Epilepsia*. 2009;50(6):1474-1483.
- Kubota BY, Coan AC, Yasuda CL, Cendes F. T2 hyperintense signal in patients with temporal lobe epilepsy with MRI signs of hippocampal sclerosis and in patients with temporal lobe epilepsy with normal MRI. *Epilepsy Behav*. 2015;46:103-108.
- Bernasconi A, Cendes F, Theodore WH, et al. Recommendations for the use of structural magnetic resonance imaging in the care of patients with epilepsy: a consensus report from the International League Against Epilepsy Neuroimaging Task Force. *Epilepsia*. 2019;60(6):1054-1068.
- Fonov V, Evans AC, Botteron K, Almli CR, McKinstry RC, Collins DL. Unbiased average age-appropriate atlases for pediatric studies. *NeuroImage*. 2011;54(1):313-327.

22. Caldaïrou B, Bernhardt BC, Kulaga-Yoskovitz J, Kim H, Bernasconi N, Bernasconi A. A surface patch-based segmentation method for hippocampal subfields. In: Ourselin S, Joskowicz L, Sabuncu MR, Unal G, Wells W, eds. *Medical Image Computing And Computer-Assisted Intervention: MICCAI 2016: Lecture Notes in Computer Science*. Springer International Publishing; 2016:379-387.
23. Kulaga-Yoskovitz J, Bernhardt BC, Hong S-J, et al. Multi-contrast submillimetric 3 Tesla hippocampal subfield segmentation protocol and dataset. *Sci Data*. 2015;2:150059.
24. Kim H, Bernhardt BC, Kulaga-Yoskovitz J, Caldaïrou B, Bernasconi A, Bernasconi N. Multivariate hippocampal subfield analysis of local MRI intensity and volume: application to temporal lobe epilepsy. In: Golland P, Hata N, Barillot C, Hornegger J, Howe R, eds. *Medical Image Computing and Computer-Assisted Intervention: MICCAI 2014*. Springer International Publishing; 2014:170-178.
25. Bernhardt BC, Bernasconi A, Liu M, et al. The spectrum of structural and functional imaging abnormalities in temporal lobe epilepsy. *Ann Neurol*. 2016;80(1):142-153.
26. Hastie T, Tibshirani R, Friedman J. Model assessment and selection. In: *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer; 2009:219-259.
27. Luo W, Phung D, Tran T, et al. Guidelines for developing and reporting machine learning predictive models in biomedical research: a multidisciplinary view. *J Med Internet Res*. 2016;18(12):e323.
28. Brier GW. Verification of forecasts expressed in terms of probability. *Mon Weather Rev*. 1950;78(1):1-3.
29. Gaizo JD, Mofrad N, Jensen JH, et al. Using machine learning to classify temporal lobe epilepsy based on diffusion MRI. *Brain Behav*. 2017;7(10):e00801.
30. Morgan VL, Sonmez Turk HH, Gore JC, Abou-Khalil B. Lateralization of temporal lobe epilepsy using resting functional magnetic resonance imaging connectivity of hippocampal networks. *Epilepsia*. 2012;53(9):1628-1635.
31. Yang Z, Choupan J, Reutens D, Hocking J. Lateralization of temporal lobe epilepsy based on resting-state functional magnetic resonance imaging and machine learning. *Front Neurol*. 2015;6:184.
32. Barron DS, Fox PT, Pardoe H, et al. Thalamic functional connectivity predicts seizure laterality in individual TLE patients: application of a biomarker development strategy. *NeuroImage Clin*. 2015;7:273-280.
33. 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.
34. Hadar PN, Kini LG, Coto C, et al. Clinical validation of automated hippocampal segmentation in temporal lobe epilepsy. *NeuroImage Clin*. 2018;20:1139-1147.
35. Wang S, Summers RM. Machine learning and radiology. *Med Image Anal*. 2012;16(5):933-951.
36. Mo J, Liu Z, Sun K, et al. Automated detection of hippocampal sclerosis using clinically empirical and radiomics features. *Epilepsia*. 2019;60(12):2519-2529.
37. Blümcke I, Arzimanoglou A, Beniczky S, Wiebe S. Roadmap for a competency-based educational curriculum in epileptology: report of the Epilepsy Education Task Force of the International League Against Epilepsy. *Epileptic Disord Int Epilepsy J Videotape*. 2019;21(2):129-140.
38. Goubiran M, Rudko DA, Santyr B, et al. In vivo normative atlas of the hippocampal subfields using multi-echo susceptibility imaging at 7 Tesla. *Hum Brain Mapp*. 2014;35(8):3588-3601.
39. Wyler AR, Dohan FC, Schweitzer JB, Berry AD. A grading system for mesial temporal pathology (hippocampal sclerosis) from anterior temporal lobectomy. *J Epilepsy*. 1992;5(4):220-225.
40. Cardinale F, Rizzi M, Vignati E, et al. Stereoelectroencephalography: retrospective analysis of 742 procedures in a single centre. *Brain*. 2019;142(9):2688-2704.