Chapter 27

Machine Learning in Neuroimaging of Epilepsy

Hyo Min Lee, Ravnoor Singh Gill, Neda Bernasconi, and Andrea Bernasconi

Abstract

Epilepsy is a prevalent chronic condition affecting about 50 million people worldwide. A third of patients suffer from seizures unresponsive to medication. Uncontrolled seizures damage the brain, are associated with cognitive decline, and have negative impact on well-being. For these patients, the surgical resection of the brain region that gives rise to seizures is the most effective treatment. In this context, due to its unmatched spatial resolution and whole-brain coverage, magnetic resonance imaging (MRI) plays a central role in detecting lesions. The last decade has witnessed an increasing use of machine learning applied to multimodal MRI, which has allowed the design of tools for computer-aided diagnosis and prognosis. In this chapter, we focus on automated algorithms for the detection of epileptogenic lesions and imaging-derived prognostic markers, including response to anti-seizure medication, postsurgical seizure outcome, and cognitive reserves. We also highlight advantages and limitations of these approaches and discuss future directions toward person-centered care.

Key words Epilepsy, Focal cortical dysplasia, Temporal lobe epilepsy

1 Introduction

Epilepsy is a prevalent chronic condition affecting about 50 million people worldwide. Seizures are generally defined as transient symptoms and signs due to excessive neuronal activity; based on these manifestations, they can be classified as focal or generalized. Various etiologies have been associated with epilepsy, including structural, genetic, infectious, metabolic, and immune. Frequent structural pathologies include traumatic brain injury, tumors, vascular malformations, stroke, and developmental disorders. A third of patients suffer from seizures unresponsive to medication [1]. Drug-resistant seizures damage the brain [2] and are associated with high risks for socioeconomic difficulties, cognitive decline, and mortality [3]. The main forms of drug-resistant focal epilepsy are related to focal cortical dysplasia (FCD), a structural brain developmental malformation, and mesiotemporal lobe sclerosis, a

Olivier Colliot (ed.), *Machine Learning for Brain Disorders*, Neuromethods, vol. 197, https://doi.org/10.1007/978-1-0716-3195-9_27,

histopathological lesion that combines various degrees of neuronal loss and gliosis in the hippocampus and adjacent cortices. To date, the most effective treatment has been the surgical resection of these structural lesions. In this context, magnetic resonance imaging (MRI) has been instrumental in the pre-surgical evaluation, as it can reliably detect these anomalies due to its unmatched spatial resolution and whole-brain coverage. Indeed, localizing a structural lesion on MRI is the strongest predictor of favorable seizure outcome after surgery [4–6]. Yet, challenges remain. Large numbers of patients have subtle lesions undetected on routine MRI. In these patients, referred to as "MRI-negative," the surgical outcome is poorer compared to those in whom a structural lesion is identified [7]. Moreover, even in carefully selected patients, about 30% may continue having seizures after surgery. These shortcomings have motivated the development of advanced analytic techniques for the discovery of diagnostic and prognostic biomarkers, which serve as input to machine learning. MRI quantitation holds promise to match or exceed the evaluation by human experts. In this chapter, we will describe algorithms for the detection of epileptogenic lesions, prediction of clinical outcomes, and identification of disease subtypes in drug-resistant focal epilepsy. We will highlight their advantages and limitations and discuss future directions toward personalized care.

2 Lesion Mapping

In epilepsy, identifying a structural lesion on MRI is crucial for successful surgery [5]. Advances in MRI acquisition technology, specifically high (3T) and ultrahigh (7T) field imaging combined with multiple phased array head coils, have permitted precise lesion characterization. Machine learning holds great promise for exceeding human performance [8]. Indeed, application on structural MRI data has enabled increasingly reliable detection of epileptogenic lesions, including those overlooked on routine radiological examination. Automated lesion detection is generally performed by supervised classifiers that are trained to learn the distributions and inter-relations between MRI features that distinguish lesional from non-lesional tissue, leveraging this knowledge to classify a given tissue type in previously unseen patients.

2.1 Mapping Hippocampal Sclerosis in Temporal Lobe Epilepsy

Temporal lobe epilepsy (TLE), the most common focal syndrome in adults, is pathologically defined by varying degrees of neuronal loss and gliosis in the hippocampus and adjacent structures [9]. On MRI, marked hippocampal sclerosis (HS) appears as atrophy and signal hyperintensity, generally more severe ipsilateral to the seizure focus. Accurate identification of hippocampal atrophy as a marker of HS is crucial for deciding the side of surgery. While volumetry

has been one of the first computational analyses applied to TLE [10–15], the need for accurate localization of pathology has motivated a move from whole-structure volumetry to surface-based approaches allowing a precise mapping of anomalies along the hippocampal axis. In this context, 3D surface-based shape models permit localizing regional morphological differences that may not be readily identifiable [16]. Surface modeling based on spherical harmonics [17] has been particularly performant [18]. Following this method, hippocampal labels are processed using a series of spherical harmonics with increasing degree of complexity to parametrize their surface boundary. Anatomical intersubject correspondence is guaranteed by aligning the surfaces of each individual to the centroid and the longitudinal axis of the first-order ellipsoid of the mean surface template derived from controls and patients. Computing the Jacobian determinants of the surface displacement vectors allows quantifying localized areas of atrophy [18, 19]. Overall, surface-based methods have proven superior to their volumetric counterparts not only in terms of segmentation performance [20] but also in predicting clinical outcomes as well as mapping disease progression [21, 22]. Applying clustering to surface-based morphometry of the hippocampus, amygdala, and entorhinal cortex, a clinically homogeneous cohort of drug-resistant TLE patients with a unilateral seizure focus could be segregated into classes with distinct MRI and histopathological signatures [23]. Extending this methodology by extracting features along the medial surface of hippocampal subfields has allowed to further probe the laminar integrity of this structure [24, 25].

Manual hippocampal volumetry is time-prohibitive and prone to rater bias. These challenges, together with increasing demand to study larger patient cohorts, have motivated the shift toward automated segmentation, setting the basis for large-scale clinical use. Initial methods for whole hippocampal segmentation used a single template or deformable models constrained by shape priors obtained from neurotypical individuals [26–29]. More recent approaches rely on multiple templates and label fusion; by selecting a subset of atlases from a template library which best fit the structure to segment, thereby accounting for intersubject variability, these approaches have provided increased performance [30–32]. In epilepsy, SurfMulti achieved identical performance in TLE (Dice: 86.9%) and healthy controls (87.5%), outperforming the widely used FreeSurfer, even in the presence of prevalent atypical hippocampal morphology (i.e., maldevelopment or malrotation) and significant atrophy [20]. Advances in MR acquisition hardware and sequence technology, which enable submillimetric resolution and improved signal-to-noise ratio, have facilitated accurate identification of hippocampal subfields or subregions, including the dentate gyrus, subiculum, and the cornu ammonis (CA1–4) regions

[33]. Several methods have been developed for MRI-based subfield segmentation [19, 34–38], providing an average Dice of 88%, with fast inference times. Among them, the SurfPatch subfield segmentation algorithm, operating on T1-weighted MRI, combines multiple templates, parametric surfaces, and patch-based sampling for compact representation of shape, texture, and intensity [38]. SurfPatch showed high segmentation accuracy (Dice >0.82 for all subfields) and robustness to the size of template library and image resolution (millimetric and sub-millimetric) while demonstrating utility for reliable TLE lateralization (93% accuracy).

Brain segmentation may serve as the basis to extract features used to train classifiers for predictions. An SVM-based classifier using volumetric features derived from whole-brain T1-weighted images was able to classify and lateralize TLE [39]. However, regions identifying TLE groups were primarily located outside the mesiotemporal lobe, making such design impractical for previously unseen cases and difficult to interpret in MRI-negative patients. Overall, while high lateralization performance (>90%) may be achieved in MRI-positive patients, the yield in MRI-negative TLE remains at less than 20% when using features derived from T1-weighted images [40, 41]. On the other hand, classifiers operating on FLAIR [42] and double inversion recovery [43] have shown 70% lateralization in MRI-negative patients. Yet, studies have been rather limited in sample size and have lacked histological verification or long-term measures of seizure outcome after surgery; moreover, absence of validation in independent datasets has precluded assessment of generalizability. To tackle these shortcomings, our group recently designed an automated surface-based linear discriminant classifier trained on T1- and FLAIR-derived laminar features of HS (Fig. 1) [44]. As HS is typically characterized by T1-weighted hypointensity and T2-weighted hyperintensity, the synthetic contrast FLAIR/T1 maximized their combined contributions to detect the full pathology spectrum. The classifier accurately lateralized the focus in 85% of patients with MRI-negative but histologically verified HS. Notably, similar high performance was achieved in two independent validation cohorts, thereby establishing generalizability across cohorts, scanners, and parameters. Such validated classifiers set the basis for broad clinical translation.

Recently, the widespread adoption of deep learning in medical imaging has promoted a resurgence in volumetric segmentation methods. Unlike contemporary algorithms, deep learning does not require the data to be extensively preprocessed, thus eliminating the need to build template libraries. More specifically, the ability of convolutional neural networks to learn salient features from multimodal data in the course of the training process rather than using hand-crafted features has enabled them to outperform

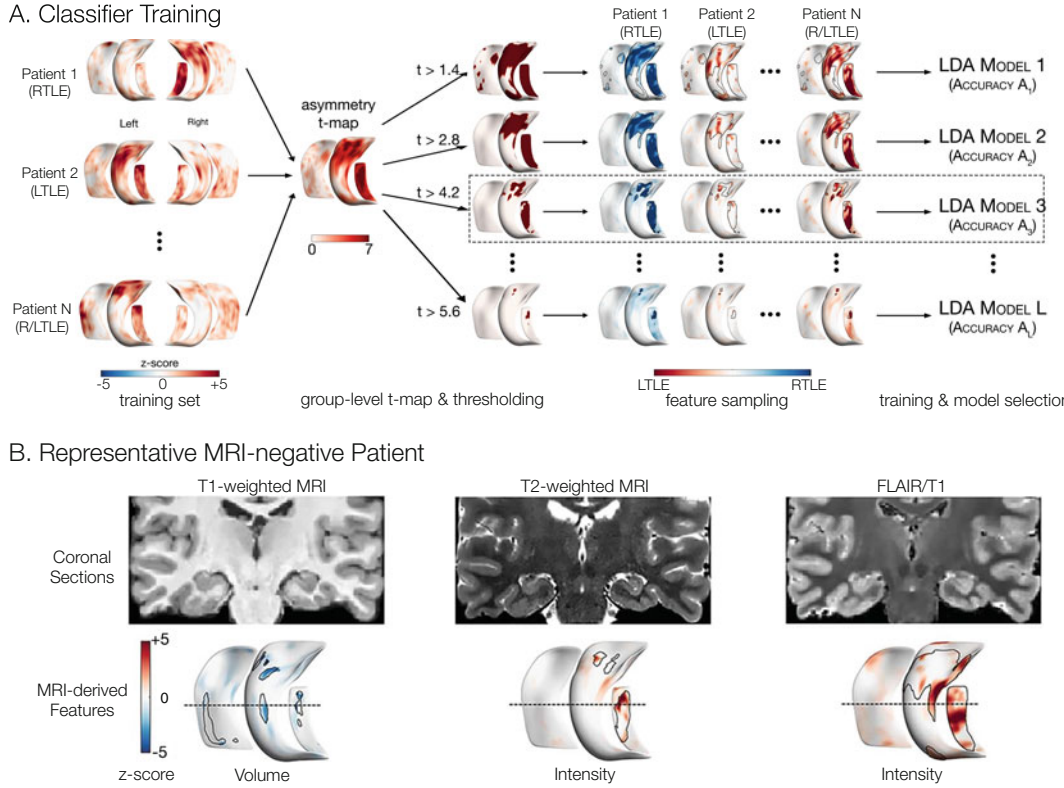

Fig. 1 Automated lateralization of hippocampal sclerosis. (**a**). In the training phase, an optimal region of interest is defined for each modality to systematically sample features (T1-derived volume, T2-weighted intensity, and FLAIR/T1 intensity) across individuals. To this purpose, in each patient paired t-tests compare corresponding vertices of the left and right subfields, z-scored with respect to healthy controls. The resulting group-level asymmetry t-map is then thresholded from 0 to the highest value and binarized; for each threshold, the binarized t-map is overlaid on the asymmetry map of each individual to compute the average across subfields. Then a linear discriminant classifier is trained for each threshold, and the model yielding the highest lateralization accuracy (in this example LDA model 3) is used to test the classifier. (**b**). Lateralization prediction in a patient with MRI-negative left TLE. Coronal sections 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 level of the coronal MRI section and the optimal ROI obtained during training is outlined in black

traditional approaches, with Dice overlap indices exceeding 90% in both healthy [44–46] and atrophic [47] hippocampi. Deep learning applications for seizure focus lateralization have insofar been limited. One study showed that deep learning classifiers performed similar or worse than SVM-based classifiers [48]; this work, however, explored only a singular set of hyperparameters using pre-defined features for the neural networks, thereby missing the opportunity to exploit hierarchical feature learning, one of the most distinctive characteristics of deep learning.

2.2 Automated Detection of Focal Cortical Dysplasia

On MRI, focal cortical dysplasia (FCD) presents with a visibility spectrum encompassing variable degrees of gray matter (GM) and white matter (WM) changes that can challenge visual identification. Indeed, recent series indicate that up to 33% of FCD Type II, the most common surgically amenable developmental malformation, present with "unremarkable" routine MRI, even though typical features are ultimately identified in the histopathology of the resected tissue [49–51]. These so-called "MRI-negative" FCDs represent a major diagnostic challenge. Indeed, to define the epileptogenic area, patients undergo long and costly hospitalizations for EEG monitoring with intracerebral electrodes, a procedure that carries risks similar to surgery itself [52, 53]. Moreover, patients without MRI evidence for FCD are less likely to undergo surgery and consistently show worse seizure control compared to those with visible lesions [4, 54, 55]. This clinical difficulty has motivated the development of computer-aided methods aimed at optimizing detection in vivo. Such techniques provide distinct information through quantitative assessment without the cost of additional scanning time.

Early methods opted for voxel-based methods to quantify group-level structural abnormalities related to MRI-visible dysplasias by thresholding GM concentration (e.g., >1 SD relative to the mean in healthy controls). While such methods are sensitive (87–100%) in detecting conspicuous malformations, they fail to identify two-thirds of subtle, MRI-negative lesions [56–59]. To counter the relative lack of specificity, our group introduced an original approach to integrate key voxel-wise textures and morphological modeling (i.e., cortical thickening, blurring of the GM-WM junction, and intensity alterations) derived from T1-weighted images into a composite map [60, 61]. The clinical value of this computer-aided visual identification was supported by its 88% sensitivity and 95% specificity, vastly outperforming conventional MRI. An alternative method quantifies blurring as voxels that belong neither to GM or WM [62]. Integrating morphological operators with higher-order image texture features invisible to the human eye into a fully automated classifier provided a sensitivity of 80% [63, 64]. In contrast to voxel-based methods, surface-based morphometry offers an anatomically plausible quantification of structural integrity that preserves cortical topology. Surface-based modeling of cortical thickness, folding complexity, and sulcal depth, together with intra- and subcortical mapping of MRI intensities and textures, allow for a more sensitive description of FCD pathology. Over the last decade, several such algorithms have been developed, with detection rates up to 83% [65–71]. The addition of FLAIR has contributed to further increase in sensitivity, particularly for the detection of smaller lesions [66]. Notably, an integration of surface-based methods into clinical workflow would be contingent to careful verification of preprocessing steps, including manual

corrections of tissue segmentation and surface extraction to obtain high-fidelity FCD features. Without such careful and intensive data preprocessing and inspection, the performance is rather poor, as demonstrated by a recent multicenter study in which the sensitivity was below 70% with a specificity close to chance level even in MRI visible lesions [72].

Despite efforts dedicated to the development of increasingly sophisticated detection algorithms, some pitfalls are to be considered. Algorithms have not been systematically validated with histologically verified lesions or independent datasets. Many have not been tested or fail in MRI-negative cases. In general, detection algorithms have assumed structural anomalies to be homogeneous across lesions and patients, a notion challenged by recent histopathological [73, 74] and genetic [75] data. Moreover, they rely on limited number of features designed by human experts based on their knowledge, which may not capture the full pathological complexity. Importantly, the deterministic nature of these algorithms does not permit risk assessment, a necessity for integration into clinical diagnostic systems. Currently, benchmark automated detection fails in 20–40% of patients, particularly those with subtle FCD, and suffers from high false-positive rates. Relative to conventional methods, in recent years, deep neural networks have shown high sensitivity at detection across various diseases [see 76, 77, for review]. Specifically, convolutional neural networks learn abstract concepts from high-dimensional data alleviating the challenging task of handcrafting features [78]. To date, a few studies have used deep learning for FCD detection [79–81]. However, their clinical description has been scarce or absent, and the information on how lesions were labeled for the training as well as histological validation was not provided. Notably, while their performance was reasonably high in MRI-positive cohorts (range: 85–92%; no MRI-negative cases identified) using either T1-weighted or T2-weighted FLAIR images, sample sizes were limited to 10–40 and sourced from a single center. Deep learning requires large corpus of expertly labeled annotations (ground truth) to train and optimize the network, both cost- and time-consuming endeavors, resulting in suboptimal cohort sizes. To overcome this challenge, our group leveraged a patch-based augmentation that extracts several hundreds of overlapping patches from a single subject, thereby scaling up the data without the requirement of an impractically large cohort [82]. This deep learning algorithm relied on clinically available T1- and T2-weighted FLAIR MRI of a large cohort of patients with histologically validated lesions, collated across multiple tertiary epilepsy centers (Fig. 2). Notably, operating on 3D voxel space (i.e., in true volumetric domain) allowed assessing the spatial neighborhood of the lesion, whereas prior surface-based methods have considered each vertex location independently. This convolutional neural network classifier yields the highest

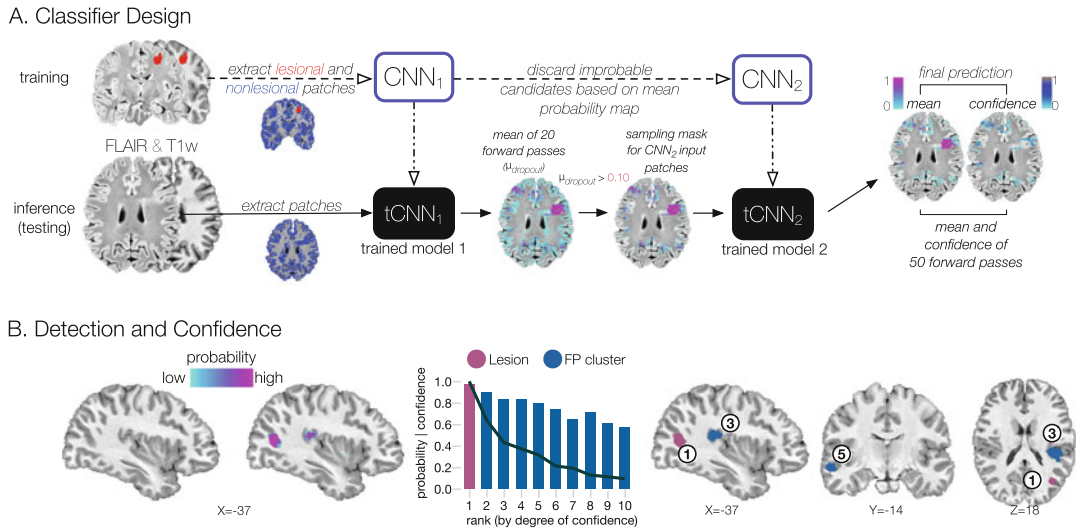

Fig. 2 Automated FCD detection using deep learning. (**a**). The training and testing workflow. In this cascaded system, the output of the convolutional neural network 1 (CNN-1) serves as an input to CNN-2. CNN-1 maximizes the detection of lesional voxels; CNN-2 reduces the number of misclassified voxels, removing false positives (FPs) while maintaining optimal sensitivity. The training procedure (dashed arrows) operating on T1-weighted and FLAIR extracts 3D patches from lesional and non-lesional tissue to yield tCNN-1 (trained model 1) and tCNN-2 with optimized weights (vertical dashed-dotted arrows). These models are then used for subject-level inference. For each unseen subject, the inference pipeline (solid arrows) uses tCNN-1 and generates a mean ($\mu_{dropout}$) of 20 predictions (forward passes); the mean map is then thresholded voxel-wise to discard improbable lesion candidates $\mu_{dropout} > 0.1$). The resulting binary mask serves to sample the input patches for the tCNN-2. A mean probability and uncertainty maps are obtained by collating 50 predictions; uncertainty is transformed into confidence. The sampling strategy (identical for training and inference) is only illustrated for testing. (**b**). Sagittal sections show the native T1-weighted MRI superimposed with the lesion probability map. The bar plot shows the probability of the lesion (purple) and false-positive (FP, blue) clusters sorted by their rank; the superimposed line indicates the degree of confidence for each cluster. In this example, the lesion (cluster 1 in purple) has both the highest probability and confidence

performance to date with a sensitivity of 93% using a leave-one-site-out cross-validation and 83% when tested on an independent cohort while maintaining a high specificity of 89% both in healthy and disease controls. Importantly, deep learning detected MRI-negative FCD with 85% sensitivity, thus offering a considerable gain over standard radiological assessment. Results were generalizable across cohorts with variable age, hardware, and sequence parameters. Using Bayesian uncertainty estimation that enables risk stratification [83, 84], our predictions were stratified according to the confidence to be truly lesional. In 73% of cases, the FCD was among the top five clusters with the highest confidence to be lesional; in half of them, it ranked the highest. Ranking putative lesional clusters in each patient based on confidence helps the examiner to gauge the significance of all findings. In other words,

by pairing predictions with risk stratification, this classifier may assist clinicians to adjust hypotheses relative to other tests, thus increasing diagnostic confidence. Taken together, such characteristics and performance promise great potential for broad clinical translation.

3 Prediction of Clinical Outcomes

While science investigating the neurobiology of epilepsy has been growing rapidly, translating knowledge into clinical practice has been limited. Specifically, individualized predictions of drug resistance, surgical outcome, and cognitive dysfunction have been attempted with limited success [85]. For example, early investigations that aimed to predict anti-seizure medication response used machine learning on genomic data (viz., single nucleotide polymorphisms) and showed limited generalizability with inconsistent performance across studies [86–88]. Similarly, other models trained on electro-clinical and demographic features of thousands of patients [89–92] achieved high sensitivity (>90%) but unacceptably low specificity (<25%). Importantly, no external validation was performed on independent cohorts. The prediction of seizure outcome after surgery has been extensively explored in TLE patients. Some of the early investigations relied on clinical [93] and neuropsychological features [94], achieving high performance, but in limited samples of less than 20 patients. Given the increasing conceptualization of TLE as a system-level disorder, numerous studies have tested the hypothesis that structural and functional alterations beyond the mesial temporal lobe may contribute to negative seizure outcome [95, 96]. For instance, WM microstructural features derived from diffusion tensor imaging have shown to achieve high sensitivity (70–86%) but modest specificity (65–70%) [97, 98]. Other studies have relied on connectivity features for prediction; these include nodal hubness of the thalamus and whole-brain distance-based measures of functional connectivity, which achieve an accuracy at about 75% but modest specificity (ranging from 35 to 62%) [99, 100]. Conversely, while topological features of structural connectome have generally shown high predictive value for favorable postsurgical outcome, with an area under the receiver operating characteristics of 0.88, specificity for prediction of seizure relapse is low (29–54%) [101, 102]. Overall, the lack of large-scale external validation and relatively low specificity of these models need to be addressed to establish their generalizability and potential clinical use.

3.1 Disease Biotyping: Leveraging Individual Variability to Optimize Predictions

To date, most neuroimaging studies of epilepsy have been based on "one-size-fits-all" group-level analytical approaches. While such study designs can isolate reliable and consistent average group-level differences, they merely decipher the common patterns without modeling the inter-individual variations along the disease spectrum [103]. Conversely, the conceptualization of epilepsy as a heterogeneous disorder and explicit modeling of inter-individual phenotypic variations may be exploited to predict individual-specific clinical outcomes [104].

Over the past decades, FCD characterization has been driven by histology, with the primary objective to establish subtype-specific imaging signatures [105]. Although histological grading is a well-defined framework, the current approach is based on descriptive criteria that do not consider the severity of each feature, thereby limiting neurobiological understanding. The ability to perform in vivo patient stratification is gaining relevance due to the emergence of minimally invasive surgical procedures that do not provide specimens for histological examination [106]. From a neurobiological standpoint, whether FCD IIB (dysmorphic neurons and balloon cells) and IIA (dysmorphic neurons only) subtypes represent etiologically distinct entities, or a spectrum is a matter of debate. Recent studies have shown significant cellular variability, with anomalies that may vary across lesions within the same subtype [73]. Moreover, multiple subtypes may coexist within the same FCD, with the most severe phenotype determining the final diagnosis [74]. Furthermore, recent studies have identified regulatory genes of the mTOR pathway that cause FCD via somatic mutations, revealing a genetic continuum not linked to discrete FCD subtypes [75]. Hence, assessing the intra- and inter-lesional variability on MRI may offer a novel basis to advance our understanding of FCD neurobiology and improve lesion detection. Leveraging hierarchical clustering to model connectivity from FCD tissue to the rest of the cortex demonstrated that network dysfunction can dissociate patients with excellent from those with suboptimal post-surgical seizure outcomes [107]. Another recent work applying consensus clustering to multi-contrast 3T MRI uncovered FCD tissue classes with distinct structural profiles, variably expressed within and across patients [108]. Importantly, these classes had differential histopathological embeddings, and their clinical utility was supported by gain in performance of a lesion detection algorithm trained on class-informed data compared to class-naïve paradigm.

In TLE, histopathological reports have shown substantial variability in the distribution and severity of mesiotemporal lobe sclerosis between patients [109, 110]. A modern approach combining quantitative histology and unsupervised machine learning identified histological subtypes with differential severity and regional signatures [111]. Motivated by these findings, recent studies have

exploited inter-individual variability of imaging or cognitive phenotypes to optimize predictions of clinical outcomes. The first attempts were based on categorical models, which provided subtypes of patients with a given phenotype. Clustering applied to surface-based morphometry uncovered four TLE subtypes having distinct subregional patterns of mesiotemporal atrophy [23]. These four subtypes differed with respect to histopathology and postsurgical seizure outcome. Classifiers operating on class membership accurately predicted surgical outcome in >90% of patients, outperforming learners trained on conventional MRI volumetry. In the context of cognition, unsupervised techniques have identified phenotypes, such as language and memory impairment associated with distinct patterns of WM microstructural damage [112] and connectome disorganization [113].

Compared to categorical models such as clustering, dimensional approaches allow a more in-depth conceptualization of inter-individual variability by uncovering axes of pathology that are co-expressed within and between individuals. In other words, such approaches allow patients to express multiple disease factors to varying degrees rather than assigning subjects to a single subtype. Applying latent Dirichlet allocation, an unsupervised technique derived from topic modeling, to multimodal MRI features of hippocampal and whole-brain GM and WM pathology, a recent study uncovered dimensions of heterogeneity (or disease factors) in TLE that were not expressed in healthy controls and only minimally in patients with frontal lobe epilepsy, supporting specificity (Figs. 3 and 4) [114]. Importantly, classifiers trained on the patients' factor composition predicted response to anti-seizure medications (76% accuracy) and surgery (88%) as well as cognitive scores for verbal IQ, memory, and sequential motor tapping, outperforming learners trained on group-level data [114]. In translational terms, assessing inter-individual variability through dimensional modeling mines clinically relevant disease characteristics that would otherwise be missed.

4 Conclusion and Future Perspectives

Machine learning applied to MRI has successfully uncovered mesoscopic structural and functional biomarkers predictive of clinical outcomes. Overall, the most significant impact has been the development of lesion detection algorithms that have transformed MRI-negative into MRI-positive, thus offering the life-changing benefits of epilepsy surgery to more patients. More recently, biotyping techniques exploiting intra- and intersubject variability have permitted to further optimize the prediction of outcomes. Integrating such approaches with other domains such as genomics

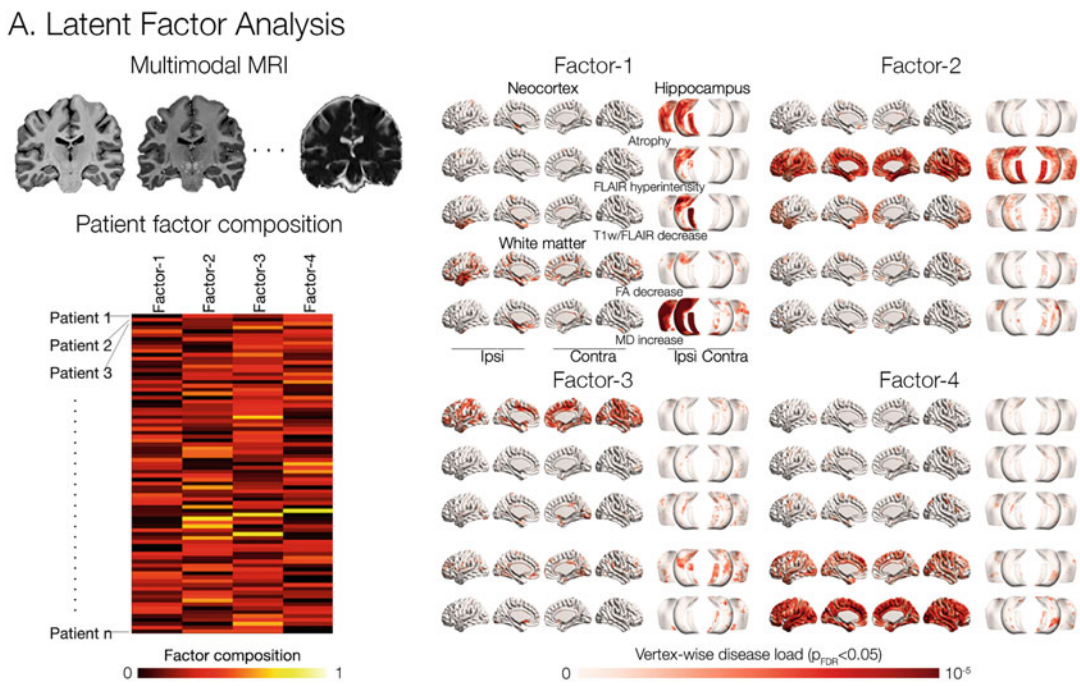

Fig. 3 Latent disease factors in TLE. Multimodal MRI (T1w, FLAIR, T1w/FLAIR, diffusion-derived FA, and MD) is combined with surface-based analysis to model the main features of TLE pathology (atrophy, gliosis, demyelination, and microstructural damage), which are z-scored with respect to the analogous vertices of healthy controls' ipsi- and contralateral to the seizure focus. Latent Dirichlet allocation uncovered four latent relations (viz., disease factors) from these features (expressed as posterior probability) and quantified their co-expression (ranging from 0 to 1) as shown in the patients' factor composition matrix. On the color scale below, the disease factor maps higher probability (darker red) and signifies a greater contribution of a given feature to the factor, namely, the disease load (pFDR <0.05)

promises to elucidate molecular mechanisms that drive MRI phenotypes, offering novel avenues to study disease processes [115, 116].

Notwithstanding its diagnostic capabilities, machine learning is still viewed by some as a "black box," possibly due to the increasing complexity of the predictive models, particularly those relying on deep learning [117]. In this regard, increased model interpretability may prevent biases and reduce the risk of incorrect clinical inferences. It is, therefore, crucial to understand how the model arrived at a particular decision. For large-scale neural networks, this may be achieved by visualizing on a map the features learned in the course of training. Besides transparency, significant obstacles to clinical adoption are privacy and ethics. These concerns have been circumvented so far through single site designs or multi-institutional training aggregating data in a single center. While the latter allows addressing model generalizability through physical access to independent datasets, federated learning may provide decentralized collaborations without data sharing [30]. As the

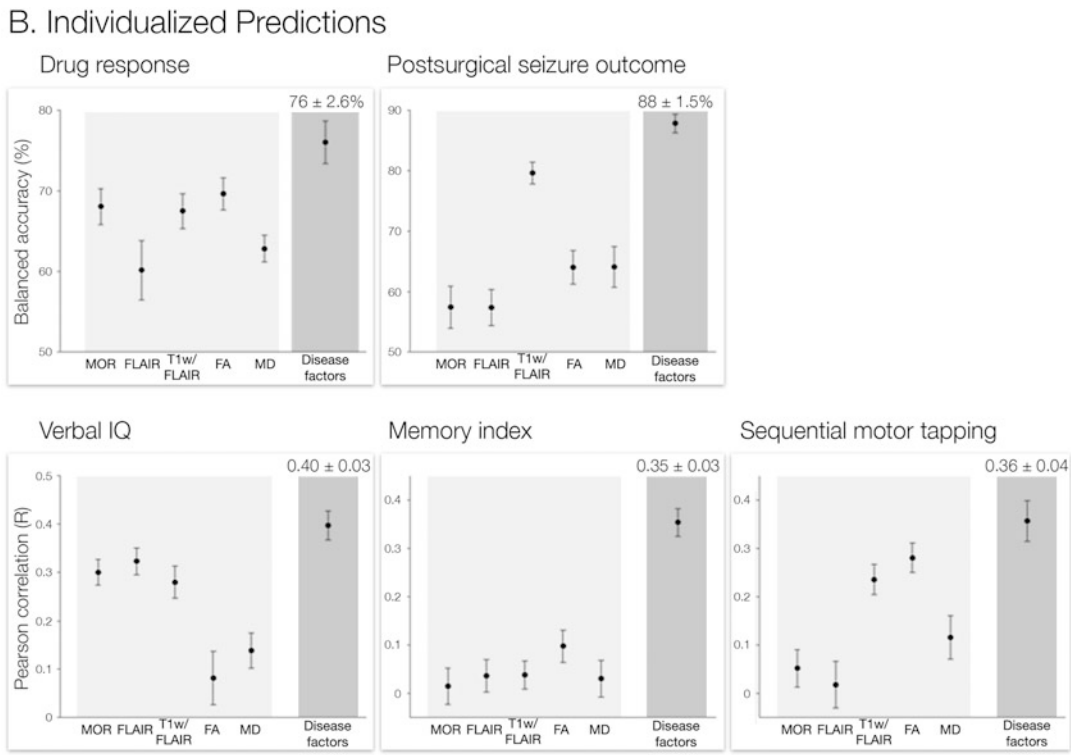

Fig. 4 Latent disease factors in TLE. Drug response, seizure outcome, verbal IQ, memory index, and motor index are more accurately predicted when using latent disease factors than when relying on conventional group-level features (pFDR <0.001). Data points indicate mean balanced accuracy for categorical data (drug-response, seizure outcome) and Pearson correlation coefficients for numerical data (cognitive scores) evaluated based on 100 repetitions of tenfold cross-validation

data corpus diversifies and expands to include more edge cases, performance and confidence of future classifiers will inevitably improve. Ultimately, clinical translation of complex techniques into practice is contingent to continued efforts in education of clinicians combined with increased accessibility to source codes and algorithms.

References

1. Wiebe S, Jette N (2012) Pharmacoresistance and the role of surgery in difficult to treat epilepsy. Nat Rev Neurol 8:669. https://doi.org/10.1038/nrneurol.2012.181
2. Caciagli L, Bernasconi A, Wiebe S, Koepp MJ, Bernasconi N, Bernhardt BC (2017) A meta-analysis on progressive atrophy in intractable temporal lobe epilepsy. Time is brain? 89(5): 506–516. https://doi.org/10.1212/wnl.0000000000004176
3. Keezer MR, Sisodiya SM, Sander JW (2016) Comorbidities of epilepsy: current concepts and future perspectives. Lancet Neurol 15(1):106–115
4. Jobst BC, Cascino GD (2015) Resective epilepsy surgery for drug-resistant focal epilepsy: a review. JAMA 313(3):285–293. https://doi.org/10.1001/jama.2014.17426
5. Téllez-Zenteno JF, Ronquillo LH, Moien-Afshari F, Wiebe S (2010) Surgical outcomes

in lesional and non-lesional epilepsy: a systematic review and meta-analysis. Epilepsy Res 89(2):310–318. https://doi.org/10.1016/j.eplepsyres.2010.02.007

6. West S, Nevitt SJ, Cotton J, Gandhi S, Weston J, Sudan A, Ramirez R, Newton R (2019) Surgery for epilepsy. Cochrane Database Syst Rev 6:CD010541
7. Bernasconi A, Bernasconi N, Bernhardt BC, Schrader D (2011) Advances in MRI for cryptogenic epilepsies. Nat Rev Neurol 7:99. https://doi.org/10.1038/nrneurol.2010.199
8. Rauschecker AM, Rudie JD, Xie L, Wang J, Duong MT, Botzolakis EJ, Kovalovich AM, Egan J, Cook TC, Bryan RN (2020) Artificial intelligence system approaching neuroradiologist-level differential diagnosis accuracy at brain MRI. Radiology 295(3): 626–637
9. Blümcke I, Thom M, Aronica E, Armstrong DD, Bartolomei F, Bernasconi A, Bernasconi N, Bien CG, Cendes F, Coras R, Cross JH, Jacques TS, Kahane P, Mathern GW, Miyata H, Moshé SL, Oz B, Özkara C, Perucca E, Sisodiya S, Wiebe S, Spreafico R (2013) International consensus classification of hippocampal sclerosis in temporal lobe epilepsy: a Task Force report from the ILAE Commission on Diagnostic Methods. Epilepsia 54(7):1315–1329. https://doi.org/doi:10.1111/epi.12220
10. Cascino GD, Jack CR Jr, Parisi JE, Sharbrough FW, Hirschorn KA, Meyer FB, Marsh WR, O'Brien PC (1991) Magnetic resonance imaging–based volume studies in temporal lobe epilepsy: pathological correlations. Ann Neurol 30(1):31–36
11. Cendes F, Andermann F, Gloor P, Evans A, Jones-Gotman M, Watson C, Melanson D, Olivier A, Peters T, Lopes-Cendes I (1993) MRI volumetric measurement of amygdala and hippocampus in temporal lobe epilepsy. Neurology 43(4):719–719
12. Watson C, Jack CR, Cendes F (1997) Volumetric magnetic resonance imaging: clinical applications and contributions to the understanding of temporal lobe epilepsy. Arch Neurol 54(12):1521–1531
13. Bernasconi N, Bernasconi A, Caramanos Z, Antel S, Andermann F, Arnold DL (2003) Mesial temporal damage in temporal lobe epilepsy: a volumetric MRI study of the hippocampus, amygdala and parahippocampal region. Brain 126(2):462–469
14. Bernasconi N, Bernasconi A, Andermann F, Dubeau F, Feindel W, Reutens D (1999) Entorhinal cortex in temporal lobe epilepsy. Quantitative MRI Study 52(9):1870–1870. https://doi.org/10.1212/wnl.52.9.1870
15. Bernasconi N, Bernasconi A, Caramanos Z, Dubeau F, Richardson J, Andermann F, Arnold D (2001) Entorhinal cortex atrophy in epilepsy patients exhibiting normal hippocampal volumes. Neurology 56(10): 1335–1339. https://doi.org/10.1212/wnl.56.10.1335
16. Hogan RE, Wang L, Bertrand ME, Willmore LJ, Bucholz RD, Nassif AS, Csernansky JG (2004) MRI-based high-dimensional hippocampal mapping in mesial temporal lobe epilepsy. Brain 127(8):1731–1740
17. Styner M, Oguz I, Xu S, Brechbühler C, Pantazis D, Levitt JJ, Shenton ME, Gerig G (2006) Framework for the statistical shape analysis of brain structures using SPHARM-PDM. Insight J 1071:242
18. Kim H, Besson P, Colliot O, Bernasconi A, Bernasconi N (2008) Surface-based vector analysis using heat equation interpolation: a new approach to quantify local hippocampal volume changes. In:Medical image computing and computer-assisted intervention—MICCAI 2008. Lecture Notes in Computer Science, vol 5241, pp 1008–1015. https://doi.org/10.1007/978-3-540-85988-8_120
19. Kim H, Bernhardt BC, Kulaga-Yoskovitz J, Caldairou B, Bernasconi A, Bernasconi N (2014) Multivariate hippocampal subfield analysis of local MRI intensity and volume: application to temporal lobe epilepsy. In: Medical image computing and computer-assisted intervention—MICCAI 2014. Lecture Notes in Computer Science, vol 8674, pp 170–178
20. Kim H, Mansi T, Bernasconi N, Bernasconi A (2012) Surface-based multi-template automated hippocampal segmentation: application to temporal lobe epilepsy. Med Image Anal 16(7):1445–1455. https://doi.org/10.1016/j.media.2012.04.008
21. Bernhardt BC, Worsley KJ, Kim H, Evans AC, Bernasconi A, Bernasconi N (2009) Longitudinal and cross-sectional analysis of atrophy in pharmacoresistant temporal lobe epilepsy. Neurology 72(20):1747–1754. https://doi.org/10.1212/01.wnl.0000345969.57574.f5
22. Bernhardt BC, Kim H, Bernasconi N (2013) Patterns of subregional mesiotemporal disease progression in temporal lobe epilepsy. Neurology 81(21):1840–1847
23. Bernhardt BC, Hong SJ, Bernasconi A, Bernasconi N (2015) Magnetic resonance imaging pattern learning in temporal lobe epilepsy: classification and prognostics. Annal Neurol

77(3):436–446. https://doi.org/10.1002/ana.24341

24. Kim H, Mansi T, Bernasconi N (2013) Disentangling hippocampal shape anomalies in epilepsy. Front Neurol 4. https://doi.org/10.3389/fneur.2013.00131
25. Bernhardt BC, Bernasconi A, Liu M, Hong SJ, Caldairou B, Goubran M, Guiot MC, Hall J, Bernasconi N (2016) The spectrum of structural and functional imaging abnormalities in temporal lobe epilepsy. Ann Neurol 80(1):142–153. https://doi.org/10.1002/ana.24691
26. Yang J, Duncan JS (2004) 3D image segmentation of deformable objects with joint shape-intensity prior models using level sets. Med Image Anal 8(3):285–294. https://doi.org/10.1016/j.media.2004.06.008
27. Pitiot A, Delingette H, Thompson PM, Ayache N (2004) Expert knowledge-guided segmentation system for brain MRI. Neuroimage 23:S85–S96. https://doi.org/10.1016/j.neuroimage.2004.07.040
28. Duchesne S, Pruessner JC, Collins DL (2002) Appearance-based segmentation of medial temporal lobe structures. Neuroimage 17(2):515–531. https://doi.org/10.1006/nimg.2002.1188
29. Khan AR, Wang L, Beg MF (2008) FreeSurfer-initiated fully-automated subcortical brain segmentation in MRI using large deformation diffeomorphic metric mapping. Neuroimage 41(3):735–746. https://doi.org/10.1016/j.neuroimage.2008.03.024
30. Wang H, Das SR, Suh JW, Altinay M, Pluta J, Craige C, Avants B, Yushkevich PA (2011) A learning-based wrapper method to correct systematic errors in automatic image segmentation: consistently improved performance in hippocampus, cortex and brain segmentation. Neuroimage 55(3):968–985. https://doi.org/10.1016/j.neuroimage.2011.01.006
31. Aljabar P, Heckemann RA, Hammers A, Hajnal JV, Rueckert D (2009) Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. Neuroimage 46(3):726–738. https://doi.org/10.1016/j.neuroimage.2009.02.018
32. Collins DL, Pruessner JC (2010) Towards accurate, automatic segmentation of the hippocampus and amygdala from MRI by augmenting ANIMAL with a template library and label fusion. Neuroimage 52(4):1355–1366. https://doi.org/10.1016/j.neuroimage.2010.04.193
33. Kulaga-Yoskovitz J, Bernhardt BC, Hong SJ, Mansi T, Liang KE, van der Kouwe AJW, Smallwood J, Bernasconi A, Bernasconi N (2015) Multi-contrast submillimetric 3 Tesla hippocampal subfield segmentation protocol and dataset. Sci Data 2:150059–150059. https://doi.org/10.1038/sdata.2015.59
34. Iglesias JE, Augustinack JC, Nguyen K, Player CM, Player A, Wright M, Roy N, Frosch MP, McKee AC, Wald LL (2015) A computational atlas of the hippocampal formation using ex vivo, ultra-high resolution MRI: application to adaptive segmentation of in vivo MRI. Neuroimage 115:117–137
35. Pipitone J, Park MTM, Winterburn J, Lett TA, Lerch JP, Pruessner JC, Lepage M, Voineskos AN, Chakravarty MM, Initiative ADN (2014) Multi-atlas segmentation of the whole hippocampus and subfields using multiple automatically generated templates. Neuroimage 101:494–512
36. Van Leemput K, Bakkour A, Benner T, Wiggins G, Wald LL, Augustinack J, Dickerson BC, Golland P, Fischl B (2009) Automated segmentation of hippocampal subfields from ultra-high resolution in vivo MRI. Hippocampus 19(6):549–557
37. Kim H, Mansi T, Bernasconi N, Bernasconi A (2011) Robust surface-based multi-template automated algorithm to segment healthy and pathological hippocampi. In: Medical image computing and computer-assisted intervention—MICCAI 2011. Lecture notes in computer science, vol 6893, pp 445–453. https://doi.org/10.1007/978-3-642-23626-6_55
38. Caldairou B, Bernhardt BC, Kulaga-Yoskovitz J, Kim H, Bernasconi N, Bernasconi A (2016) A surface patch-based segmentation method for hippocampal subfields. In: International conference on medical image computing and computer-assisted intervention—MICCAI 2016. Lecture notes in computer science, vol 9901, pp 379–387. https://doi.org/10.1007/978-3-319-46723-8_44
39. Keihaninejad S, Heckemann RA, Gousias IS, Hajnal JV, Duncan JS, Aljabar P, Rueckert D, Hammers A (2012) Classification and lateralization of temporal lobe epilepsies with and without hippocampal atrophy based on whole-brain automatic MRI segmentation. PLoS One 7(4):e33096
40. Hadar PN, Kini LG, Coto C, Piskin V, Callans LE, Chen SH, Stein JM, Das SR, Yushkevich PA, Davis KA (2018) Clinical validation of

automated hippocampal segmentation in temporal lobe epilepsy. Neuroimage Clin 20: 1139–1147

41. Mahmoudi F, Elisevich K, Bagher-Ebadian H, Nazem-Zadeh MR, Davoodi-Bojd E, Schwalb JM, Kaur M, Soltanian-Zadeh H (2018) Data mining MR image features of select structures for lateralization of mesial temporal lobe epilepsy. PLoS One 13(8): e0199137
42. Beheshti I, Sone D, Maikusa N, Kimura Y, Shigemoto Y, Sato N, Matsuda H (2020) FLAIR-wise machine-learning classification and lateralization of MRI-negative 18F-FDG PET-positive temporal lobe epilepsy. Front Neurol 11(1433). https://doi.org/10.3389/fneur.2020.580713
43. Beheshti I, Sone D, Maikusa N, Kimura Y, Shigemoto Y, Sato N, Matsuda H (2021) Accurate lateralization and classification of MRI-negative 18F-FDG-PET-positive temporal lobe epilepsy using double inversion recovery and machine-learning. Comput Biol Med 137:104805. https://doi.org/10.1016/j.compbiomed.2021.104805
44. Caldairou B, Foit NA, Mutti C, Fadaie F, Gill R, Lee HM, Demerath T, Urbach H, Schulze-Bonhage A, Bernasconi A (2021) An MRI-based machine learning prediction framework to lateralize hippocampal sclerosis in patients with temporal lobe epilepsy. Neurology 97(16):e1583–e1593
45. Manjon JV, Romero JE, Coupe P (2020) DeepHIPS: a novel deep learning based hippocampus subfield segmentation method. arXiv preprint arXiv:200111789
46. Zhu H, Shi F, Wang L, Hung SC, Chen MH, Wang S, Lin W, Shen D (2019) Dilated dense U-net for infant hippocampus subfield segmentation. Front Neuroinform 13(30). https://doi.org/10.3389/fninf.2019.00030
47. Goubran M, Ntiri EE, Akhavein H, Holmes M, Nestor S, Ramirez J, Adamo S, Ozzoude M, Scott C, Gao F, Martel A, Swardfager W, Masellis M, Swartz R, MacIntosh B, Black SE (2020) Hippocampal segmentation for brains with extensive atrophy using three-dimensional convolutional neural networks. Hum Brain Mapp 41(2): 291–308. https://doi.org/10.1002/hbm.24811
48. Gleichgerrcht E, Munsell BC, Alhusaini S, Alvim MK, Bargalló N, Bender B, Bernasconi A, Bernasconi N, Bernhardt B, Blackmon K (2021) Artificial intelligence for classification of temporal lobe epilepsy with ROI-level MRI data: a worldwide ENIGMA-Epilepsy study. NeuroImage Clin 31:102765
49. Cohen-Gadol AA, Özduman K, Bronen RA, Kim JH, Spencer DD (2004) Long-term outcome after epilepsy surgery for focal cortical dysplasia. J Neurosurg 101(1):55–65. https://doi.org/10.3171/jns.2004.101.1.0055
50. Krsek P, Maton B, Korman B, Pacheco-Jacome E, Jayakar P, Dunoyer C, Rey G, Morrison G, Ragheb J, Vinters HV, Resnick T, Duchowny M (2008) Different features of histopathological subtypes of pediatric focal cortical dysplasia. Ann Neurol 63(6):758–769. https://doi.org/10.1002/ana.21398
51. Krsek P, Maton B, Jayakar P, Dean P, Korman B, Rey G, Dunoyer C, Pacheco-Jacome E, Morrison G, Ragheb J, Vinters HV, Resnick T, Duchowny M (2009) Incomplete resection of focal cortical dysplasia is the main predictor of poor postsurgical outcome. Neurology 72(3):217–223. https://doi.org/10.1212/01.wnl.0000334365.22854.d3
52. Widdess-Walsh P, Jeha L, Nair D, Kotagal P, Bingaman W, Najm I (2007) Subdural electrode analysis in focal cortical dysplasia: predictors of surgical outcome. Neurology 69(7):660–667
53. Hedegärd E, Bjellvi J, Edelvik A, Rydenhag B, Flink R, Malmgren K (2014) Complications to invasive epilepsy surgery workup with subdural and depth electrodes: a prospective population-based observational study. J Neurol Neurosurg Psychiatry 85(7):716–720
54. Fauser S, Schulze-Bonhage A, Honegger J, Carmona H, Huppertz HJ, Pantazis G, Rona S, Bast T, Strobl K, Steinhoff BJ, Korinthenberg R, Rating D, Volk B, Zentner J (2004) Focal cortical dysplasias: surgical outcome in 67 patients in relation to histological subtypes and dual pathology. Brain 127(11):2406–2418. https://doi.org/10.1093/brain/awh277
55. Cascino GD (2004) Surgical treatment for epilepsy. Epilepsy Res 60(2):179–186. https://doi.org/10.1016/j.eplepsyres.2004.07.003
56. Focke NK, Symms MR, Burdett JL, Duncan JS (2008) Voxel-based analysis of whole brain FLAIR at 3T detects focal cortical dysplasia. Epilepsia 49(5):786–793
57. Rugg-Gunn F, Eriksson S, Boulby P, Symms M, Barker G, Duncan J (2003)

Magnetization transfer imaging in focal epilepsy. Neurology 60(10):1638–1645

58. Rugg-Gunn F, Boulby P, Symms M, Barker G, Duncan J (2005) Whole-brain T2 mapping demonstrates occult abnormalities in focal epilepsy. Neurology 64(2):318–325
59. Salmenpera TM, Symms MR, Rugg-Gunn FJ, Boulby PA, Free SL, Barker GJ, Yousry TA, Duncan JS (2007) Evaluation of quantitative magnetic resonance imaging contrasts in MRI-negative refractory focal epilepsy. Epilepsia 48(2):229–237
60. Bernasconi A, Antel SB, Collins DL, Bernasconi N, Olivier A, Dubeau F, Pike GB, Andermann F, Arnold DL (2001) Texture analysis and morphological processing of magnetic resonance imaging assist detection of focal cortical dysplasia in extra-temporal partial epilepsy. Ann Neurol 49(6):770–775. https://doi.org/10.1002/ana.1013
61. Colliot O, Antel SB, Naessens VB, Bernasconi N, Bernasconi A (2006) In vivo profiling of focal cortical dysplasia on high-resolution MRI with computational models. Epilepsia 47(1):134–142. https://doi.org/10.1111/j.1528-1167.2006.00379.x
62. Huppertz HJ, Grimm C, Fauser S, Kassubek J, Mader I, Hochmuth A, Spreer J, Schulze-Bonhage A (2005) Enhanced visualization of blurred gray–white matter junctions in focal cortical dysplasia by voxel-based 3D MRI analysis. Epilepsy Res 67(1–2):35–50
63. Antel SB, Li LM, Cendes F, Collins DL, Kearney RE, Shinghal R, Arnold DL (2002) Predicting surgical outcome in temporal lobe epilepsy patients using MRI and MRSI. Neurology 58(10):1505–1512. https://doi.org/10.1212/wnl.58.10.1505
64. Antel SB, Collins DL, Bernasconi N, Andermann F, Shinghal R, Kearney RE, Arnold DL, Bernasconi A (2003) Automated detection of focal cortical dysplasia lesions using computational models of their MRI characteristics and texture analysis. Neuroimage 19(4):1748–1759. https://doi.org/10.1016/S1053-8119(03)00226-X
65. Adler S, Wagstyl K, Gunny R, Ronan L, Carmichael D, Cross JH, Fletcher PC, Baldeweg T (2016) Novel surface features for automated detection of focal cortical dysplasias in paediatric epilepsy. Neuroimage Clin 14:18–27. https://doi.org/10.1016/j.nicl.2016.12.030
66. Gill RS, Hong SJ, Fadaie F, Caldairou B, Bernhardt B, Bernasconi N, Bernasconi A (2017) Automated detection of epileptogenic cortical malformations using multimodal MRI. In: Deep learning in medical image analysis and multimodal learning for clinical decision support: third international workshop, DLMIA 2017, ML-CDS 2017. Lecture notes in computer science, vol 10553, pp 349–356. https://doi.org/10.1007/978-3-319-67558-9_40
67. Hong SJ, Kim H, Schrader D, Bernasconi N, Bernhardt B, Bernasconi A (2014) Automated detection of cortical dysplasia type II in MRI-negative epilepsy. Neurology 83(1):48–55. https://doi.org/10.1212/WNL.0000000000000543
68. Jin B, Krishnan B, Adler S, Wagstyl K, Hu W, Jones S, Najm I, Alexopoulos A, Zhang K, Zhang J (2018) Automated detection of focal cortical dysplasia type II with surface-based magnetic resonance imaging postprocessing and machine learning. Epilepsia 59(5):982–992
69. Tan YL, Kim H, Lee S, Tihan T, Ver Hoef L, Mueller SG, Barkovich AJ, Xu D, Knowlton R (2018) Quantitative surface analysis of combined MRI and PET enhances detection of focal cortical dysplasias. Neuroimage 166:10–18
70. Snyder K, Whitehead EP, Theodore WH, Zaghloul KA, Inati SJ, Inati SK (2021) Distinguishing type II focal cortical dysplasias from normal cortex: a novel normative modeling approach. Neuroimage Clin 30:102565
71. Kini LG, Gee JC, Litt B (2016) Computational analysis in epilepsy neuroimaging: a survey of features and methods. Neuroimage Clin 11:515–529
72. Spitzer H, Ripart M, Whitaker K, Napolitano A, De Palma L, De Benedictis A, Foldes S, Humphreys Z, Zhang K, Hu W, Mo J, Likeman M, Davies S, Guttler C, Lenge M, Cohen NT, Tang Y, Wang S, Chari A, Tisdall M, Bargallo N, Conde-Blanco E, Pariente JC, Pascual-Diaz S, Delgado-Martínez I, Pérez-Enríquez C, Lagorio I, Abela E, Mullatti N, O'Muircheartaigh J, Vecchiato K, Liu Y, Caligiuri M, Sinclair B, Vivash L, Willard A, Kandasamy J, McLellan A, Sokol D, Semmelroch M, Kloster A, Opheim G, Ribeiro L, Yasuda C, Rossi-Espagnet C, Zhang K, Hamandi K, Tietze A, Barba C, Guerrini R, Gaillard WD, You X, Wang I, González-Ortiz S, Severino M, Striano P, Tortora D, Kalviainen R, Gambardella A, Labate A, Desmond P, Lui E, O'Brien T, Shetty J, Jackson G, Duncan J, Winston G, Pinborg L, Cendes F, Theis FJ, Shinohara RT, Cross JH, Baldeweg T, Adler S, Wagstyl K (2021) Interpretable surface-based detection of focal cortical dysplasias: a MELD study.

medRxiv. https://doi.org/10.1101/2021.12.13.21267721
73. Najm IM, Sarnat HB, Blümcke I (2018) Review: the international consensus classification of Focal Cortical Dysplasia—a critical update 2018. Neuropathol Appl Neurobiol 44(1):18–31. https://doi.org/10.1111/nan.12462
74. Iffland PH, Crino PB (2017) Focal cortical dysplasia: gene mutations, cell signaling, and therapeutic implications. Annu Rev Pathol 12(1):547–571. https://doi.org/10.1146/annurev-pathol-052016-100138
75. Marsan E, Baulac S (2018) Review: mechanistic target of rapamycin (mTOR) pathway, focal cortical dysplasia and epilepsy. Neuropathol Appl Neurobiol 44(1):6–17. https://doi.org/10.1111/nan.12463
76. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, Van Der Laak JA, Van Ginneken B, Sánchez CI (2017) A survey on deep learning in medical image analysis. Med Image Anal 42:60–88
77. Topol EJ (2019) High-performance medicine: the convergence of human and artificial intelligence. Nat Med 25(1):44–56
78. Goodfellow I, Bengio Y, Courville A (2016) Deep learning. MIT Press, Cambridge, MA
79. Dev KB, Jogi PS, Niyas S, Vinayagamani S, Kesavadas C, Rajan J (2019) Automatic detection and localization of Focal Cortical Dysplasia lesions in MRI using fully convolutional neural network. Biomed Signal Process Control 52:218–225
80. Thomas E, Pawan S, Kumar S, Horo A, Niyas S, Vinayagamani S, Kesavadas C, Rajan J (2020) Multi-res-attention UNet: a CNN model for the segmentation of focal cortical dysplasia lesions from magnetic resonance images. IEEE J Biomed Health Inf 25(5):1724–1734
81. Wang H, Ahmed SN, Mandal M (2020) Automated detection of focal cortical dysplasia using a deep convolutional neural network. Comput Med Imaging Graph 79:101662
82. Gill RS, Lee HM, Caldairou B, Hong SJ, Barba C, Deleo F, D'Incerti L, Coelho VCM, Lenge M, Semmelroch M (2021) Multicenter validation of a deep learning detection algorithm for focal cortical dysplasia. Neurology 97(16):e1571–e1582
83. Leibig C, Allken V, Ayhan MS, Berens P, Wahl S (2017) Leveraging uncertainty information from deep neural networks for disease detection. Sci Rep 7(1):17816. https://doi.org/10.1038/s41598-017-17876-z
84. Gal Y, Ghahramani Z (2015) Bayesian convolutional neural networks with Bernoulli approximate variational inference. arXiv preprint arXiv:150602158
85. Smolyansky ED, Hakeem H, Ge Z, Chen Z, Kwan P (2021) Machine learning models for decision support in epilepsy management: a critical review. Epilepsy Behav 123:108273
86. Petrovski S, Szoeke CE, Sheffield LJ, D'souza W, Huggins RM, O'brien TJ (2009) Multi-SNP pharmacogenomic classifier is superior to single-SNP models for predicting drug outcome in complex diseases. Pharmacogenet Genomics 19(2):147–152
87. Shazadi K, Petrovski S, Roten A, Miller H, Huggins RM, Brodie MJ, Pirmohamed M, Johnson MR, Marson AG, O'Brien TJ (2014) Validation of a multigenic model to predict seizure control in newly treated epilepsy. Epilepsy Res 108(10):1797–1805
88. Silva-Alves MS, Secolin R, Carvalho BS, Yasuda CL, Bilevicius E, Alvim MK, Santos RO, Maurer-Morelli CV, Cendes F, Lopes-Cendes I (2017) A prediction algorithm for drug response in patients with mesial temporal lobe epilepsy based on clinical and genetic information. PLoS One 12(1):e0169214
89. An S, Malhotra K, Dilley C, Han-Burgess E, Valdez JN, Robertson J, Clark C, Westover MB, Sun J (2018) Predicting drug-resistant epilepsy—a machine learning approach based on administrative claims data. Epilepsy Behav 89:118–125
90. Devinsky O, Dilley C, Ozery-Flato M, Aharonov R, Goldschmidt Y, Rosen-Zvi M, Clark C, Fritz P (2016) Changing the approach to treatment choice in epilepsy using big data. Epilepsy Behav 56:32–37
91. Delen D, Davazdahemami B, Eryarsoy E, Tomak L, Valluru A (2020) Using predictive analytics to identify drug-resistant epilepsy patients. Health Inf J 26(1):449–460
92. Yao L, Cai M, Chen Y, Shen C, Shi L, Guo Y (2019) Prediction of antiepileptic drug treatment outcomes of patients with newly diagnosed epilepsy by machine learning. Epilepsy Behav 96:92–97
93. Memarian N, Kim S, Dewar S, Engel Jr J, Staba RJ (2015) Multimodal data and machine learning for surgery outcome prediction in complicated cases of mesial temporal lobe epilepsy. Comput Biol Med 64:67–78
94. Armañanzas R, Alonso-Nanclares L, DeFelipe-Oroquieta J, Kastanauskaite A, de Sola RG, DeFelipe J, Bielza C, Larrañaga P (2013) Machine learning approach for the

outcome prediction of temporal lobe epilepsy surgery. PLoS One 8(4):e62819

95. Bernhardt B, Hong SJ, Bernasconi A, Bernasconi N (2013) Imaging structural and functional brain networks in temporal lobe epilepsy. Front Hum Neurosci 7(624). https://doi.org/10.3389/fnhum.2013.00624
96. Caciagli L, Bernhardt BC, Hong SJ, Bernasconi A, Bernasconi N (2014) Functional network alterations and their structural substrate in drug-resistant epilepsy. Front Neurosci 8:411
97. Munsell BC, Wee CY, Keller SS, Weber B, Elger C, da Silva LAT, Nesland T, Styner M, Shen D, Bonilha L (2015) Evaluation of machine learning algorithms for treatment outcome prediction in patients with epilepsy based on structural connectome data. Neuroimage 118:219–230
98. Taylor PN, Sinha N, Wang Y, Vos SB, De Tisi J, Miserocchi A, McEvoy AW, Winston GP, Duncan JS (2018) The impact of epilepsy surgery on the structural connectome and its relation to outcome. Neuroimage Clin 18:202–214
99. He X, Doucet GE, Pustina D, Sperling MR, Sharan AD, Tracy JI (2017) Presurgical thalamic hubness predicts surgical outcome in temporal lobe epilepsy. Neurology 88(24):2285–2293
100. Larivière S, Weng Y, Vos de Wael R, Royer J, Frauscher B, Wang Z, Bernasconi A, Bernasconi N, Schrader DV, Zhang Z, Bernhardt BC (2020) Functional connectome contractions in temporal lobe epilepsy: microstructural underpinnings and predictors of surgical outcome. Epilepsia 61(6):1221–1233. https://doi.org/10.1111/epi.16540
101. Gleichgerrcht E, Keller SS, Drane DL, Munsell BC, Davis KA, Kaestner E, Weber B, Krantz S, Vandergrift WA, Edwards JC (2020) Temporal lobe epilepsy surgical outcomes can be inferred based on structural connectome hubs: a machine learning study. Ann Neurol 88(5):970–983
102. Sinha N, Wang Y, da Silva NM, Miserocchi A, McEvoy AW, de Tisi J, Vos SB, Winston GP, Duncan JS, Taylor PN (2021) Structural brain network abnormalities and the probability of seizure recurrence after epilepsy surgery. Neurology 96(5):e758–e771
103. Lo A, Chernoff H, Zheng T, Lo SH (2015) Why significant variables aren't automatically good predictors. Proc Natl Acad Sci USA 112(45):13892–13897. https://doi.org/10.1073/pnas.1518285112
104. Arbabshirani MR, Plis S, Sui J, Calhoun VD (2017) Single subject prediction of brain disorders in neuroimaging: promises and pitfalls. Neuroimage 145:137–165. https://doi.org/10.1016/j.neuroimage.2016.02.079
105. 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(10):1065–1077. https://doi.org/10.1007/s00234-012-1049-1
106. Gross RE, Stern MA, Willie JT, Fasano RE, Saindane AM, Soares BP, Pedersen NP, Drane DL (2018) Stereotactic laser amygdalohippocampotomy for mesial temporal lobe epilepsy. Ann Neurol 83(3):575–587. https://doi.org/10.1002/ana.25180
107. Hong SJ, Lee HM, Gill R, Crane J, Sziklas V, Bernhardt BC, Bernasconi N, Bernasconi A (2019) A connectome-based mechanistic model of focal cortical dysplasia. Brain 142(3):688–699. https://doi.org/10.1093/brain/awz009
108. Lee HM, Gill RS, Fadaie F, Cho KH, Guiot MC, Hong SJ, Bernasconi N, Bernasconi A (2020) Unsupervised machine learning reveals lesional variability in focal cortical dysplasia at mesoscopic scale. Neuroimage Clin 28:102438. https://doi.org/10.1016/j.nicl.2020.102438
109. Margerison J, Corsellis J (1966) Epilepsy and the temporal lobes: a clinical, electroencephalographic and neuropathologic study of the brain in epilepsy, with particular reference to the temporal lobes. Brain 89(3):499–530. https://doi.org/10.1093/brain/89.3.499
110. De Lanerolle NC, Kim JH, Williamson A, Spencer SS, Zaveri HP, Eid T, Spencer DD (2003) A retrospective analysis of hippocampal pathology in human temporal lobe epilepsy: evidence for distinctive patient subcategories. Epilepsia 44(5):677–687
111. Blümcke I, Pauli E, Clusmann H, Schramm J, Becker A, Elger C, Merschhemke M, Meencke HJ, Lehmann T, von Deimling A (2007) A new clinico-pathological classification system for mesial temporal sclerosis. Acta Neuropathol 113(3):235–244
112. Reyes A, Kaestner E, Bahrami N, Balachandra A, Hegde M, Paul BM, Hermann B, McDonald CR (2019) Cognitive phenotypes in temporal lobe epilepsy are associated with distinct patterns of white matter network abnormalities. Neurology 92(17):e1957–e1968. https://doi.org/10.1212/wnl.0000000000007370

113. Rodríguez-Cruces R, Bernhardt BC, Concha L (2020) Multidimensional associations between cognition and connectome organization in temporal lobe epilepsy. Neuroimage 213:116706. https://doi.org/10.1016/j.neuroimage.2020.116706
114. Lee HM, Fadaie F, Gill R, Caldairou B, Sziklas V, Crane J, Hong SJ, Bernhardt BC, Bernasconi A, Bernasconi N (2021) Decomposing MRI phenotypic heterogeneity in epilepsy: a step towards personalized classification. Brain 145(3):897–908. https://doi.org/10.1093/brain/awab425
115. Arnatkevičiūtė A, Fulcher BD, Fornito A (2019) A practical guide to linking brain-wide gene expression and neuroimaging data. Neuroimage 189:353–367. https://doi.org/10.1016/j.neuroimage.2019.01.011
116. Markello RD, Arnatkeviciute A, Poline JB, Fulcher BD, Fornito A, Misic B (2021) Standardizing workflows in imaging transcriptomics with the abagen toolbox. Elife 10: e72129
117. Lipton ZC (2018) The mythos of model interpretability: in machine learning, the concept of interpretability is both important and slippery. Queue 16(3):31–57