Epilepsy Currents
2022, Vol. 22(5) 279-281

DOI: 10.1177_15357597221105139
journals.sagepub.com/home/epi
SAGE

Sage Against the Machine: Promise and Challenge of Artificial Intelligence in Epilepsy

Ioannis Karakis, MD, PhD, MSc[1]

Multicenter Validation of a Deep Learning Detection Algorithm for Focal Cortical Dysplasia

Gill RS, Lee HM, Caldairou B, et al. *Neurology*. 2021 Oct 19;97(16):e1571-e1582. doi:10.1212/WNL.0000000000012698. Epub 2021 Sep 14. PMID: 34521691; PMCID: PMC8548962.

Background and Objective: To test the hypothesis that a multicenter-validated computer deep learning algorithm detects MRI-negative focal cortical dysplasia (FCD). Methods: We used clinically acquired 3-dimensional (3D) T1-weighted and 3D fluid-attenuated inversion recovery MRI of 148 patients (median age 23 years [range 2-55 years]; 47% female) with histologically verified FCD at 9 centers to train a deep convolutional neural network (CNN) classifier. Images were initially deemed MRI-negative in 51% of patients, in whom intracranial EEG determined the focus. For risk stratification, the CNN incorporated bayesian uncertainty estimation as a measure of confidence. To evaluate performance, detection maps were compared to expert FCD manual labels. Sensitivity was tested in an independent cohort of 23 cases with FCD (13 $\pm$ 10 years). Applying the algorithm to 42 healthy controls and 89 controls with temporal lobe epilepsy disease tested specificity. Results: Overall sensitivity was 93% (137 of 148 FCD detected) using a leave-one-site-out cross-validation, with an average of 6 false positives per patient. Sensitivity in MRI-negative FCD was 85%. In 73% of patients, the FCD was among the clusters with the highest confidence; in half, it ranked the highest. Sensitivity in the independent cohort was 83% (19 of 23; average of 5 false positives per patient). Specificity was 89% in healthy and disease controls. Discussion: This first multicenter-validated deep learning detection algorithm yields the highest sensitivity to date in MRI-negative FCD. By pairing predictions with risk stratification, this classifier may assist clinicians in adjusting hypotheses relative to other tests, increasing diagnostic confidence. Moreover, generalizability across age and MRI hardware makes this approach ideal for presurgical evaluation of MRI-negative epilepsy. Classification of evidence: This study provides Class III evidence that deep learning on multimodal MRI accurately identifies FCD in patients with epilepsy initially diagnosed as MRI negative.

Commentary

If you are reading this commentary on your portable electronic device, chances are that you are operating a face or voice recognition application that uses artificial intelligence (AI) to access it and that your browser uses embodied AI technology to retrieve it. If, on the other hand, you are a Gutenbergian enthusiast patiently waiting for the *Epilepsy Currents* printed version to arrive at your doorstep, the editorial office likely deployed some form of AI to identify you as a consumer and the postman utilized an AI based navigation system to deliver it to you. In other words, AI is already widely adopted in every corner of our daily lives,[1] even if it is not overtly apparent. Any field of medicine, including epilepsy, could not be an exception to this rule.

The current manuscript[2] exemplifies the use of AI in the realm of neuroimaging of epilepsy. By coalescing data from 9 tertiary epilepsy centers worldwide, the investigators created a repository of 148 children and adults with histologically-proven, type II focal cortical dysplasia (FCD), half of whom were deemed to be non-lesional on visual analysis of their pre-surgical magnetic resonance imaging (MRI). Using the gray matter volume in their high resolution (3T), three-dimensional (3D), T1-weighted and fluid-attenuated inversion recovery (FLAIR) sequences as the input, they trained an algorithm that could facilitate detection of the FCDs through deep learning. By retrospectively applying this algorithm

[1] Department of Neurology, Emory University School of Medicine, Atlanta, GA, USA

to the MRI-negative patients, they demonstrated sensitivity of 85% in detecting FCDs. Applying the algorithm on an independent cohort of 42 healthy controls and 89 controls with temporal lobe epilepsy (TLE) due to radiological and/or histologically proven hippocampal sclerosis (HS), they demonstrated a specificity of 89%.

Given that patients with MRI-negative, drug-resistant epilepsy (DRE) constitute a formidable challenge,[3] an automated detection algorithm for FCDs, one of the most commonly encountered culprits, holds significant potential to transform clinical practice. This is not the first, and it is certainly not going to be the last of a series of studies utilizing AI to detect potentially epileptogenic lesions in MRI-negative patients. Compared to prior attempts, this study utilized 3D images rather than surface-based, two-dimensional methods. It also implemented a probabilistic strategy in detecting FCDs rather than the most-commonly deployed dichotomous approach. Its main advantage though lies in the large representation of well-curated, clinical, radiological and histological data of FCD-positive patients across the age spectrum and around the globe. The technical aspects of the created repository (eg segmented lesions co-registration, detection map performance evaluation) were performed meticulously by experts. A systematic approach in controlling for potential sources of bias (eg incorporation bias, spectrum bias, risk of overfitting, generalizability threats) and drawing inferences was deployed.[3]

Despite the high quality data and rigorous methodological design leading to an overall sensitivity of 93% in the whole cohort with the leave-one-site-out cross-validation, an average of 6 false positives per patient were observed. This would not be a negligible time commitment for an expert neuroradiologist to scrutinize and, most importantly, not a trivial risk for a patient who may undergo additional electrodes implantation. Interestingly, lesions in the insula or the parahippocampal gyrus demonstrated the highest rates of false positives, perhaps accounting for the cytoarchitectonic similarities of these cortices with FCD histopathologic traits.[2] Some of the false positives, particularly in the frontocentral lesions, constituted perilesional anomalies. Given the optimal surgical outcome in many of these cases, it is evident that this algorithm may detect histologically abnormal but clinically insignificant MRI-negative regions, highlighting the importance of additional paraclinical tests and, foremost, clinical acumen in utilizing this information. One of these tests, for example, would be MR spectroscopy for differentiating between FCDs from low-grade tumors that may mimic them but could be automatically detected by the proposed algorithm.[2]

Looking at the false negatives, 6 of the 11 unresolved MRI-negative FCDs were situated in the orbitofrontal cortex, likely as a result of insufficient representation of FCDs in these regions in the training set, corroborating the need of high quality, large volume data in any AI paradigm. Interestingly, in 5 of the 11 undetected MRI-negative cases, the lesion could be identified through modulation of the probability threshold by incorporating seizure semiology and electrophysiology, attesting again to the importance of clinical reasoning and integration of any automated detection method with human oversight. Interestingly, the FCD detection sensitivity in the controlled group of TLE patients was similar between MRI-positive HS and MRI-negative HS cases, suggesting good specificity in patients with alternative lesions.[2]

Besides FCD detection, the 2 main sub-fields of AI, namely machine learning (ie, data driven statistical modeling dependent on human interventions) and deep learning (ie, application of multiple layers of neural networks with limited human intervention), have been used in many other domains of imaging epilepsy and its comorbidities.[4] Distinguishing individuals with epileptic and psychogenic non epileptic seizures from healthy controls, classifying focal and generalized epilepsies, lateralizing and localizing the epileptogenic zone, detecting and predicting seizures and neurobehavioral disturbances, gauging response to antiseizure medications, dietary treatments, neuromodulation and postoperative seizure outcome are a few only examples.[4-6] The majority of these studies used structural, diffusion and functional MRI or nuclear medicine techniques (eg positron emission tomography or single-photon emission computed tomography), but are typically fraught by small sample size, emphasis on adult patients with TLE, arduous engineering processing, limited reproducibility and poor generalizability. Beyond biomedical imaging, AI and computational approaches have been used in epilepsy to analyze clinical, neurophysiologic (eg electromyographic kinetic data, transcranial magnetic stimulation/magnetoencephalographic or scalp/intracranial video-electroencephalographic signals) and genomic/proteomic data for similar purposes.[7-9] The same limitations apply to these studies compared to their neuroimaging counterparts.

In an era when physician shortages and burnout menaces are looming, what prevents AI from being fully integrated in the practice of epilepsy? And what can be done to materialize its hype and hope? Beyond the logistical aspects that require close interdisciplinary collaboration between health care providers, computer scientists and engineers,[7] accumulating large and diverse clinical datasets, and ensuring centralized data quality assurance is key in creating accurate and highly reproducible, externally validated algorithms.[9] The focus of any algorithm should be on clinically meaningful questions benchmarked against a widely accepted "ground truth" both for model input and output.[10] In parallel, the principles of human autonomy, transparency, accountability, inclusivity and sustainability need to be taken into account.[1] Ethical and regulatory framework is paramount to avert the risk of medical negligence, prevent complacency that can threaten further innovation and ensure empathetic implementation. For any application to be widely adopted, it has to be freely available, user-friendly, prompt, readily incorporated into the workflow in real-time, and easily understood by its stakeholders, the patients and health care providers alike.

At the end of the day, the best use of artificial intelligence passes through human intelligence. In the "dilemma" of sage

against the machine, it is evident that sage with the machine would be the indomitable answer. And as it always happens with major breakthroughs in human history, it will be the inventor and operator who will define the destiny of the tool as a blessing or a curse.

ORCID iD

Ioannis Karakis https://orcid.org/0000-0001-5122-7211

References

1. Buruk B, Ekmekci PE, Arda B. A critical perspective on guidelines for responsible and trustworthy artificial intelligence. *Med Healthc Philos*. 2020 Sep;23(3):387-399. PMID: 32236794. doi:10.1007/s11019-020-09948-1
2. Gill RS, Lee H-M, Caldairou B, et al. Multicenter validation of a deep learning detection algorithm for focal cortical dysplasia. *Neurology*. 2021 Oct 19;97(16):e1571-e1582. PMID: 34521691. PMCI: PMC8548962. 10.1212/WNL.0000000000012698
3. Sinha N, Davis KA. Mapping epileptogenic tissues in MRI-negative focal epilepsy. *Neurology*. 2021 Oct 19;97(16):754-755. Epub 2021 Sep 14. PMID: 34521690. doi:10.1212/WNL.0000000000012696
4. Cendes F, McDonald CR. Artificial intelligence applications in the imaging of epilepsy and its comorbidities: Present and future. *Epilepsy Current*. 2022 Jan 12;22(2):91-96. PMID: 35444507, PMCID: PMC8988724. doi:10.1177/15357597211068600
5. Sone D, Beheshti I. Clinical application of machine learning models for brain imaging in epilepsy: A review. *Front Neurosci*. 2021 Jun 22;15:684825. PMID: 34239413. PMCID:PMC8258163. doi:10.3389/fnins.2021.684825
6. Yuan J, Ran X, Liu K, et al. Machine learning applications on neuroimaging for diagnosis and prognosis of epilepsy: A review. *J Neurosci Methods*. 2022 Feb 15;368:109441. Epub 2021 Dec 21. PMID: 34942271. doi:10.1016/j.jneumeth.2021.109441
7. An S, Kang C, Lee HW. Artificial intelligence and computational approaches for epilepsy. *J Epile Res*. 2020 Jun 30;10(1):8-17. PMID: 32983950; PMCID: PMC7494883. doi:10.14581/jer.20003
8. Kaur T, Diwakar A, Kirandeep, et al. Artificial intelligence in epilepsy. *Neurol India*. 2021 May-Jun;69(3):560-566. PMID: 34169842. doi:10.4103/0028-3886.317233
9. Abbasi B, Goldenholz DM. Machine learning applications in epilepsy. *Epilepsia*. 2019 Oct;60(10):2037-2047. Epub 2019 Sep 3. PMID: 31478577. doi:10.1111/epi.16333
10. Josephson CB, Wiebe S. Precision medicine: Academic dreaming or clinical reality? *Epilepsia*. 2021 Mar;62(suppl 2):S78-S89. Epub 2020 Nov 17. PMID: 33205406. doi:10.1111/epi.16739