Research Articles   Research Administration

7-27-2021

Artificial intelligence for classification of temporal lobe epilepsy with ROI-level MRI data: A worldwide ENIGMA-Epilepsy study

Ezequiel Gleichgerrcht

Brent C. Munsell

Saud Alhusaini

Marina K.M. Alvim

Núria Bargalló

See next page for additional authors

Authors

Ezequiel Gleichgerrcht, Brent C. Munsell, Saud Alhusaini, Marina K.M. Alvim, Núria Bargalló, Benjamin Bender, Andrea Bernasconi, Neda Bernasconi, Boris Bernhardt, Karen Blackmon, Maria Eugenia Caligiuri, Fernando Cendes, Luis Concha, Patricia M. Desmond, Orrin Devinsky, Colin P. Doherty, Martin Domin, John S. Duncan, Niels K. Focke, Antonio Gambardella, Bo Gong, Renzo Guerrini, Sean N. Hatton, Reetta Kälviäinen, Simon S. Keller, Peter Kochunov, Raviteja Kotikalapudi, Barbara A.K. Kreilkamp, Angelo Labate, Soenke Langner, Sara Larivière, Matteo Lenge, Elaine Lui, Pascal Martin, Mario Mascalchi, Stefano Meletti, Terence J. O'Brien, Heath R. Pardoe, Jose C. Pariente, Jun Xian Rao, Mark P. Richardson, Raúl Rodríguez-Cruces, Theodor Rüber, Ben Sinclair, Hamid Soltanian-Zadeh, Dan J. Stein, Pasquale Striano, Peter N. Taylor, Rhys H. Thomas, Anna Elisabetta Vaudano, Lucy Vivash, Felix von Podewills, Sjoerd B. Vos, Bernd Weber, Yi Yao, Clarissa Lin Yasuda, Junsong Zhang, Paul M. Thompson, Sanjay M. Sisodiya, Carrie R. McDonald, and Leonardo Bonilha

Artificial intelligence for classification of temporal lobe epilepsy with ROI-level MRI data: A worldwide ENIGMA-Epilepsy study

Ezequiel Gleichgerrcht [a, *], Brent C. Munsell [b,c], Saud Alhusaini [d,e], Marina K.M. Alvim [f], Núria Bargalló [g,h], Benjamin Bender [i], Andrea Bernasconi [j], Neda Bernasconi [j], Boris Bernhardt [k], Karen Blackmon [l], Maria Eugenia Caligiuri [m], Fernando Cendes [f], Luis Concha [n], Patricia M. Desmond [o], Orrin Devinsky [p], Colin P. Doherty [q,r], Martin Domin [s], John S. Duncan [t], Niels K. Focke [u], Antonio Gambardella [m,v], Bo Gong [w], Renzo Guerrini [x], Sean N. Hatton [y], Reetta Kälviäinen [z,aa], Simon S. Keller [ab,ac], Peter Kochunov [ad], Raviteja Kotikalapudi [i,ae,af], Barbara A.K. Kreilkamp [u,ab], Angelo Labate [m,v], Soenke Langner [ag,ah], Sara Larivière [k], Matteo Lenge [ai,aj], Elaine Lui [o], Pascal Martin [af], Mario Mascalchi [ak], Stefano Meletti [al,am], Terence J. O'Brien [an,ao,ap], Heath R. Pardoe [p], Jose C. Pariente [g], Jun Xian Rao [aq], Mark P. Richardson [ar], Raúl Rodríguez-Cruces [n,as], Theodor Rüber [at], Ben Sinclair [an,ao,ap], Hamid Soltanian-Zadeh [au,av], Dan J. Stein [aw], Pasquale Striano [ax,ay], Peter N. Taylor [ay,az], Rhys H. Thomas [ba], Anna Elisabetta Vaudano [al,am], Lucy Vivash [an,ao,ap], Felix von Podewills [bb], Sjoerd B. Vos [bc,bd], Bernd Weber [be], Yi Yao [be], Clarissa Lin Yasuda [f], Junsong Zhang [bf], Paul M. Thompson [bg], Sanjay M. Sisodiya [bh,bi], Carrie R. McDonald [aq], Leonardo Bonilha [a], for theENIGMA-Epilepsy Working Group

[a] *Department of Neurology, Medical University of South Carolina, Charleston, SC, USA*
[b] *Department of Psychiatry, University of North Carolina at Chapel Hill, NC, USA*
[c] *Department of Computer Science, University of North Carolina at Chapel Hill, NC, USA*
[d] *Neurology Department, Yale University School of Medicine, New Haven, CT, USA*
[e] *Department of Molecular and Cellular Therapeutics, The Royal College of Surgeons in Ireland, Dublin, Ireland*
[f] *Department of Neurology and Neuroimaging Laboratory, University of Campinas - UNICAMP, Campinas, SP, Brazil*
[g] *Magnetic Resonance Image Core Facility, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Universitat de Barcelona, Barcelona, Spain*
[h] *Department of Radiology of Center of Image Diagnosis (CDIC), Hospital Clinic de Barcelona, Barcelona, Spain*
[i] *Department of Diagnostic and Interventional Neuroradiology, University Hospital Tübingen, Tübingen, Germany*
[j] *Neuroimaging of Epilepsy Laboratory, Montreal Neurological Institute, McGill University, Montreal, QC, Canada*
[k] *McConnell Brain Imaging Center, Montreal Neurological Institute, McGill University, Montreal, QC, Canada*
[l] *Psychiatry and Psychology, Mayo Clinic, Jacksonville, FL, USA*
[m] *Neuroscience Research Center, Department of Medical and Surgical Sciences, University "Magna Græcia" of Catanzaro, Catanzaro, Italy*
[n] *Instituto de Neurobiología, Universidad Nacional Autónoma de México, Querétaro, Mexico*
[o] *Department of Radiology, Royal Melbourne Hospital, University of Melbourne, Melbourne, VIC, Australia*
[p] *Department of Neurology, Langone School of Medicine, New York University, New York, NY, USA*
[q] *Trinity College Dublin, School of Medicine, Dublin, Ireland*
[r] *FutureNeuro SFI Research Centre for Rare and Chronic Neurological Diseases, Dublin, Ireland*
[s] *Functional Imaging Unit, Department of Diagnostic Radiology and Neuroradiology, University Medicine Greifswald, Greifswald, Germany*
[t] *Department of Clinical and Experimental Epilepsy, UCL Queen Square Institute of Neurology, London, UK*
[u] *University Medicine Göttingen, Clinical Neurophysiology, Göttingen, Germany*
[v] *Institute of Neurology, University "Magna Græcia" of Catanzaro, Catanzaro, Italy*
[w] *Department of Radiology, BC Children's Hospital, University of British Columbia, Vancouver, BC, Canada*
[x] *Neuroscience Department, University of Florence, Florence, Italy*
[y] *Center for Multimodal Imaging and Genetics, University of California, San Diego, La Jolla, CA, USA*
[z] *Kuopio University Hospital, Member of EpiCARE ERN, Kuopio, Finland*
[aa] *Institute of Clinical Medicine, Neurology, University of Eastern Finland, Kuopio, Finland*
[ab] *Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool, UK*

[ac] *The Walton Centre NHS Foundation Trust, Liverpool, UK*
[ad] *Department of Psychiatry, University of Maryland School of Medicine, Baltimore, MD, USA*
[ae] *Department of Clinical Neurophysiology, University Hospital Göttingen, Goettingen, Germany*
[af] *Department of Neurology and Epileptology, Hertie Institute for Clinical Brain Research, University Hospital Tübingen, Tübingen, Germany*
[ag] *Institute for Diagnostic Radiology and Neuroradiology, University Medicine Greifswald, Greifswald, Germany*
[ah] *Institute for Diagnostic and Interventional Radiology, Pediatric and Neuroradiology, University Medical Centre Rostock, Rostock, Germany*
[ai] *Pediatric Neurology, Neurogenetics and Neurobiology Unit and Laboratories, Children's Hospital A. Meyer-University of Florence, Florence, Italy*
[aj] *Functional and Epilepsy Neurosurgery Unit, Neurosurgery Department, Children's Hospital A. Meyer-University of Florence, Florence, Italy*
[ak] *'Mario Serio' Department of Clinical and Experimental Medica Sciences, University of Florence, Florence, Italy*
[al] *Department of Biomedical, Metabolic, and Neural Sciences, University of Modena and Reggio Emilia, Modena, Italy*
[am] *Neurology Unit, OCB Hospital, AOU Modena, Modena, Italy*
[an] *Department of Neuroscience, Monash University, Melbourne, VIC, Australia*
[ao] *The Department of Medicine (The Royal Melbourne Hospital), The University of Melbourne, Parkville, VIC, Australia*
[ap] *Department of Neurology, Alfred Health, Melbourne, VIC, Australia*
[aq] *Department of Psychiatry, University of California San Diego, La Jolla, CA, USA*
[ar] *Division of Neuroscience, King's College London, London, UK*
[as] *Montreal Neurological Institute and Hospital, McGill University, Montreal, QC, Canada*
[at] *Department of Epileptology, University Hospital Bonn, Bonn, Germany*
[au] *Radiology and Research Administration, Henry Ford Health System, Detroit, MI, USA*
[av] *School of Electrical and Computer Engineering, College of Engineering, University of Tehran, Tehran, Iran*
[aw] *SA MRC Unit on Risk & Resilience in Mental Disorders, Department of Psychiatry & Neuroscience Institute, University of Cape Town, Cape Town, South Africa*
[ax] *IRCCS Istituto 'G. Gaslini', Genova, Italy*
[ay] *Department of Neurosciences, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health, University of Genova, Genova, Italy*
[az] *School of Computing, Newcastle University, Newcastle Upon Tyne, UK*
[ba] *Institute of Translational and Clinical Research, Newcastle University, Newcastle Upon Tyne, UK*
[bb] *Department of Neurology, Epilepsy Center, University Medicine Greifswald, Greifswald, Germany*
[bc] *Centre for Medical Image Computing, Department of Computer Science, University College London, London, UK*
[bd] *Neuroradiological Academic Unit, UCL Queen Square Institute of Neurology, University College London, London, UK*
[be] *Institute of Experimental Epileptology and Cognition Research, University of Bonn, Bonn, Germany*
[bf] *Cognitive Science Department, School of Informatics, Xiamen University, Xiamen, China*
[bg] *Imaging Genetics Center, Mark and Mary Stevens Institute for Neuroimaging and Informatics, Keck School of Medicine, University of Southern California, Marina del Rey, CA, USA*
[bh] *UCL Queen Square Institute of Neurology, London, UK*
[bi] *Chalfont Centre for Epilepsy, Bucks, UK*

A R T I C L E   I N F O

Keywords:
Epilepsy
Temporal lobe epilepsy
Machine learning
Artificial inteligence

A B S T R A C T

Artificial intelligence has recently gained popularity across different medical fields to aid in the detection of diseases based on pathology samples or medical imaging findings. Brain magnetic resonance imaging (MRI) is a key assessment tool for patients with temporal lobe epilepsy (TLE). The role of machine learning and artificial intelligence to increase detection of brain abnormalities in TLE remains inconclusive. We used support vector machine (SV) and deep learning (DL) models based on region of interest (ROI-based) structural (n = 336) and diffusion (n = 863) brain MRI data from patients with TLE with ("lesional") and without ("non-lesional") radiographic features suggestive of underlying hippocampal sclerosis from the multinational (multi-center) ENIGMA-Epilepsy consortium. Our data showed that models to identify TLE performed better or similar (68–75%) compared to models to lateralize the side of TLE (56–73%, except structural-based) based on diffusion data with the opposite pattern seen for structural data (67–75% to diagnose vs. 83% to lateralize). In other aspects, structural and diffusion-based models showed similar classification accuracies. Our classification models for patients with hippocampal sclerosis were more accurate (68–76%) than models that stratified non-lesional patients (53–62%). Overall, SV and DL models performed similarly with several instances in which SV mildly outperformed DL. We discuss the relative performance of these models with ROI-level data and the implications for future applications of machine learning and artificial intelligence in epilepsy care.

1. Introduction

Temporal lobe epilepsy (TLE) is the most common focal onset epilepsy in adults (Tellez-Zenteno and Hernandez-Ronquillo, 2012). Although hippocampal sclerosis is a common pathological finding in TLE (Babb and Brown, 1987; Blumcke et al., 2017), structural abnormalities in TLE are not restricted to the medial temporal region and may extend to the neocortex (Blanc et al., 2011; Margerison and Corsellis, 1966). Magnetic resonance imaging (MRI) provides a non-invasive approach to study epilepsy-related pathology, and quantitative analyses of MRI data have shown widespread grey and white matter abnormalities beyond the hippocampal structure (Whelan et al., 2018). Tissue volume abnormalities occur both proximal and distal to medial temporal lobe circuits with histopathologically confirmed hippocampal sclerosis and EEG confirmation of concordant seizure laterality (Ahmadi et al., 2009; Bernhardt et al., 2010; Bernhardt et al., 2008; Bonilha et al., 2010a; Bonilha et al., 2010b; Bonilha et al., 2003; Caligiuri et al., 2016; Focke et al., 2008; Lin et al., 2007; McDonald et al., 2008b).

Extra-hippocampal network structural abnormalities are associated with cognitive deficits (McDonald et al., 2008a; Rodriguez-Cruces et al., 2020; Yogarajah et al., 2008) as well as treatment response phenotypes (Bernhardt et al., 2015; Bonilha et al., 2015; Bonilha and Keller, 2015; Gleichgerrcht et al., 2018; Munsell et al., 2015). Methods that can accurately detect, quantify, and profile structural abnormalities at an individual level can be relevant for multiple aspects of translational epilepsy research and may help guide clinical management in the future (Gleichgerrcht and Bonilha, 2017; Gleichgerrcht et al., 2015).

Artificial intelligence offers a promising approach to assess subtle abnormalities in neuroimaging data. In particular, machine learning and deep learning may enable the detection of patterns in an automated way that would otherwise be difficult to discern through qualitative evaluations. Moreover, training and evaluation of these models generally follow a framework that emphasizes out-of-sample prediction performance, where a trained model is tested in an independent dataset to

determine whether conclusions from one sample may be generalizable and not due to biases from a small sampled population. Artificial intelligence has gained significant popularity in medicine as it can achieve detection accuracies similar to, or sometimes better than, human experts in the diagnosis of many diseases, such as lung cancer (Coudray et al., 2018), skin cancer (Haenssle et al., 2018), and retinal abnormalities (Gulshan et al., 2016). More recently, applications of artificial intelligence and machine learning have expanded in the field of epilepsy (Abbasi and Goldenholz, 2019; Focke et al., 2012; Gleichgerrcht et al., 2020; Gleichgerrcht et al., 2018; Rudie et al., 2015).

Can machine learning and deep learning accurately classify individual brain images into a pattern typical of TLE, or even lateralize TLE? This question remains unanswered, partly due to the smaller sample size of TLE datasets compared with other diseases (e.g., pneumonia, skin cancer, retinal abnormalities). The present study uses a novel initiative, the *Enhancing NeuroImaging Genetics through Meta-Analysis Epilepsy project* (ENIGMA-Epilepsy). ENIGMA-Epilepsy is a global initiative, which pools individually collected samples from sites across the world with uniform image processing and phenotypic characterizations (Thompson et al., 2020). ENIGMA-Epilepsy offers the unique advantage of sample size: it constituted the first collection of quantitative imaging data in common epilepsy syndromes, including persons with TLE, with more than one thousand data points derived from neuroimaging studies of patients as well as healthy controls with structural and/or diffusion imaging available, together with basic clinical and socio-demographic phenotypes (Hatton et al., 2020; Lariviere et al., 2020; Whelan et al., 2018). However, in its current form, it has the inherent limitation of containing only region of interest (ROI)-based data, i.e. data collected at the level of a "brain region" as defined by an atlas with predefined anatomical boundaries. This is an important trade-off: the large sample size compensates for the lower resolution of the data (i.e., data summarized at the *regional* level), although efforts to obtain raw image-based analysis are underway.

In this study, we investigated if artificial intelligence applied to ENIGMA ROI-wise data could be used to 1) classify individuals as either having TLE or being healthy controls and 2) lateralize into right versus left TLE. These are important steps in the assessment of persons with epilepsy as this can have consequences not only for their medical but also their potential surgical management if they were to remain refractory to anti-seizure medications. Contrary to prior studies ENIGMA-epilepsy which focused on ROI-level discrimination of groups, this study sought to predict group classification based on multivariate machine learning approaches. To achieve these goals, we used morphological features from T1-weighted data as well as diffusion MRI parameters sensitive to microstructural variations. We hypothesized that ML could be a viable tool for TLE classification and lateralization. We further extended our investigation by comparing the classification accuracy of different ML methods. If the hypothesis was to be confirmed, the study could serve as a guide for future multi-center large cohort initiatives.

2. Materials and methods

2.1. Participant inclusion

This retrospective study included data from adult TLE patients (aged from 18 to 70), identified using the International League Against Epilepsy classification criteria (Berg et al., 2010). Diagnoses were based on the standard of care assessment batteries at each site; this included neurophysiology and neuroimaging studies as needed on an individual basis. Based on these data, each patient was clinically classified a priori by the medical team at each site into one of four groups: (i) left-sided TLE with MRI signs of HS (TLE-HS-L), (ii) right-sided TLE with HS (TLE-HS-R), (iii) non-sclerosis left-sided TLE (TLE-NS-L), and (iv) non-sclerosis right-sided TLE (TLE-NS-R). Patients with TLE-NS were defined based on the absence of any abnormalities (including evidence of underlying hippocampal sclerosis) as deemed by each site based on

conventional radiographic interpretation. Because structural MRI data collection for ENIGMA-Epilepsy initially classified patients in the TLE-NS group as "Other," there were no structural data for TLE-NS as it is not possible to tease apart this subgroup from the remainder of the "Other" category members. Hence, only diffusion data are available for TLE-NS, as this label was implemented during the second wave of data collection. Patients with discordant imaging/neurophysiology data (e.g., left-sided onset seizures with neuroimaging findings suggestive of right-sided mesial temporal sclerosis) were not included except if histopathological confirmation allowed definitive classification. We excluded patients with progressive diseases (e.g., Rasmussen's encephalitis), autoimmune etiology, malformations of cortical development, tumors, or previous neurosurgical intervention.

Healthy controls from each site had no prior history of psychiatric or neurological disorders. The Institutional Review Board (IRB) approval for pseudo-anonymized data collection and sharing was obtained individually at each site prior to enrollment into the consortium.

2.2. Structural imaging dataset

Data from 16 sites were included in the structural MRI analysis, totaling 336 patients with TLE (187 TLE-HS-L and 149 TLE-HS-R) and 631 research center-matched healthy controls. The locations, dates, and periods of participant recruitment as well as demographic and clinical characteristics of this sample are detailed in Whelan et al. (2018). The data analyzed for this study was derived from the pipeline implemented for the original structural MRI analysis of this cohort (Whelan et al., 2018) which essentially contemplated an analysis at each site using FreeSurfer 5.3.0, for automated analysis of brain structure (Fischl, 2012). Scanning details are provided in **Supplementary Table 1**. Using the Desikan-Killiany atlas, cortical thickness and surface area measures were extracted for 34 left-hemispheric grey matter (GM) regions of interest (ROI) and 34 right-hemispheric GM ROIs. In addition, subcortical volume measures were extracted for 8 left-hemispheric GM ROIs and 8 right-hemispheric GM ROIs (152 ROI features in total - i.e., 68 thickness measures, 68 surface area measures, and 16 subcortical volume measures). Visual inspections of cortical segmentations were conducted blinded to participants' diagnosis following standardized ENIGMA protocols (available at http://enigma.ini.usc.edu) used in prior genetic studies of brain structure (Adams et al., 2016; Hibar et al., 2017; Hibar et al., 2015; Stein et al., 2012) and large-scale case-control studies of neuropsychiatric illnesses and neurological disorders (Boedhoe et al., 2017; Hoogman et al., 2017; Schmaal et al., 2017). Each site identified outliers by running a series of standardized bash scripts to detect participants with cortical thickness measures greater or less than 1.5 times the interquartile range. Outlier data were then visually inspected by overlaying the participant's cortical segmentations on their whole-brain anatomical images. Structures were excluded if judged as inaccurately segmented by the local analyst. For this study, only participants with a full dataset (i.e., no omitted volumes) were analyzed.

2.3. Diffusion data set

Data from 21 sites were included in the diffusion imaging study, totaling 863 patients (320 TLE-HS-L, 268 TLE-HS-R, 162 TLE-NS-L, and 113 TLE-NS-R) and research center-matched healthy controls (1,069 with available fractional anisotropy (FA) data, 976 with available radial diffusivity (RD) data). The locations, dates, and periods of participant recruitment as well as demographic and clinical characteristics of this sample are reported in Hatton et al. (Hatton et al., 2020).

Scanner descriptions and acquisition protocols for all sites are provided in **Supplementary Table 2.** Each site conducted the preprocessing of diffusion-weighted images, including eddy current correction, echo-planar imaging (EPI)-induced distortion correction, and tensor estimation. Next, diffusion-tensor imaging (DTI) images were processed using the ENIGMA-DTI protocols. These image processing and quality

control protocols are freely available at the ENIGMA-DTI[1] and NITRC[2] websites. For this study, we present fractional anisotropy (FA) and radial diffusivity (RD) metrics as these have recently shown the largest effect size differences across different epilepsy syndromes in a prior DTI mega-analysis by the ENIGMA-Epilepsy consortium (Hatton et al., 2020). Measures of FA and RD were hence derived for ROIs based on the Johns Hopkins University (JHU) atlas (Faria et al., 2010) which included left- and right-sided tracts as well as the midline structures of the body (BCC), genu (GCC), and splenium (SCC) of the corpus callosum (41 ROI features in total). For this study, only patients with a full dataset (i.e., no omitted volumes) were analyzed in each modality, although the exclusion of participants with at least one outlier value was modality-specific; that is, if a participant had an outlier FA value in a given ROI, that patient was still eligible for inclusion if the ROI value was a non-outlier in RD.

2.4. Multi-site data harmonization

The batch-effect correction tool "ComBat" was employed to harmonize between-site and between-protocol variations in structural and diffusion metrics, as previously done for other multi-site cohorts (Fortin et al., 2017; Hatton et al., 2020; Villalon-Reina et al., 2019; Zavaliangos-Petropulu et al., 2019). *ComBat* rescales the data for each scanner instance using a z-score transformation map common to all features using an empirical Bayes framework (Johnson et al., 2007) to adjust for variability across sites, assuming that all ROIs share the same common distribution. Prior to pipeline construction (Section 2.6), multi-site ROI, age, and sex differences were minimized using the ComBat technique.

2.5. Group imbalance correction

While the number of participants in both the structural and diffusion datasets was large, the number of healthy controls included was disproportionate relative to the number of patients with TLE, creating an imbalance problem that typically results in a classification model that is biased or overfitted to the samples in the majority group (i.e., HC in this study). To mitigate imbalance overfitting concerns, prior to pipeline construction (Section 2.6), we randomly selected ROI dataset samples (structural or diffusion) in the majority group to equal the number of minority (i.e., TLE) ROI dataset samples. This size-matching control cohort was randomly chosen from the larger pool of control participants at each of the 1000 iterations described below, maximizing selection by chance in each instance. We nonetheless acknowledge that there are alternatives approaches to prevent overfitting involving weighted penalties to the cost function of the classification algorithm based on the empirical class distribution which may achieve similar goals without the exclusion/inclusion of specific subsets of participants.

2.6. Configurable classification pipeline

A configurable pipeline allows to exchange one model approach for another (e.g. support vector machine or deep learning) without creating ad hoc properties for each framework. In general, our pipeline sequentially applied two different models to identify structural or diffusion ROI group (e.g., HC vs. TLE-HS-L) differences at the subject level (Supplementary Figure 1). To apply our classification pipeline, a subject's harmonized (Section 2.4) structural (Section 2.2) or diffusion (Section 2.3) ROI data was input into the pipeline, after which the data were sequentially processed by an ROI variable scaling model (Section 2.6.1) followed by a group classification model (Section 2.6.2), and the output was the subject's assigned group (e.g., HC, TLE-HS-L, etc.).

To consistently evaluate classification pipeline performance, while mitigating biases related to software or computing differences, three

strategies were taken: 1) a configurable pipeline approach (**Supplementary** Fig. 1B) was used that allowed individual models in the pipeline to be replaced with minimal coding changes, 2) the configurable pipeline was developed using well-known and well-tested publicly available Python software libraries (https://github.com/brent-munsell/enigma_roi), and 3) all the configurable pipeline results were generated on the same high-performance compute cluster (https://its.unc.edu/research-computing/longleaf-cluster/).

2.6.1. ROI variable scaling

As the segmentation atlases employed ROIs of different volumes, and ROI measurements may be scaled differently (i.e. ROI volume measurements have larger values than ROI surface area or cortical thickness measurements) it was important to mitigate overfitting caused by scale differences. In other words, if the volume of ROI *A* is less than that of ROI *B*, the group classification model may be overfitted (or biased) to ROI *B* measurements solely due to size. To control for this, a simple minimum and maximum scaling technique was applied to each ROI: for each column (ROI) in the input matrix, 1) the minimum value across all participants was computed, 2) the minimum value was subtracted from each participant's ROI value, 3) the maximum value was computed across all participants, and 4) each participant's value was then divided by the maximum value.

2.6.2. Group classification

Two different ML algorithms were used in this study to classify individuals into one of two groups. Specifically, we employed two supervised ML techniques to perform group classification: a support vector classifier (SVC) and a deep-learning classifier (DLC). The SVC used a linear kernel and the optimal regularization parameter (C) was determined by a grid search procedure. In general, the DLC was a fully connected neural network that applied linear activation (Supplementary Figure 2) functions and featured: (i) one input layer where the number of nodes equaled the number of ROIs (152 for structural, 41 for diffusion), (ii) two hidden layers with the optimal number of nodes in each hidden layer determined by a grid search procedure (Section 2.6), and (iii) one output layer with one node representing healthy controls and one node representing one of the TLE patient groups (**Supplementary Note 1**). Notably, the optimal DLC learning and drop-out rates were also determined by a grid search.

2.6.3. Grid search parameters

The following value ranges were used for grid search of optimal model parameters. For SVC, the C values went from 0.1 to 2.0 at 0.05 increments. For DLC, learning rate = 0.001 to 0.01 at 0.0005 increments; number of epochs = 100 to 800 at 50 increments; hidden layer 1 (L1) units = 5 to 100 at 5 increments; hidden layer 2 (L2) units = 3 to 20 at 1 increment; l_2 regularization penalty = 0.1 to 0.7 at 0.1 increments; dropout rate = 0.2 to 0.5 at 0.1 increments.
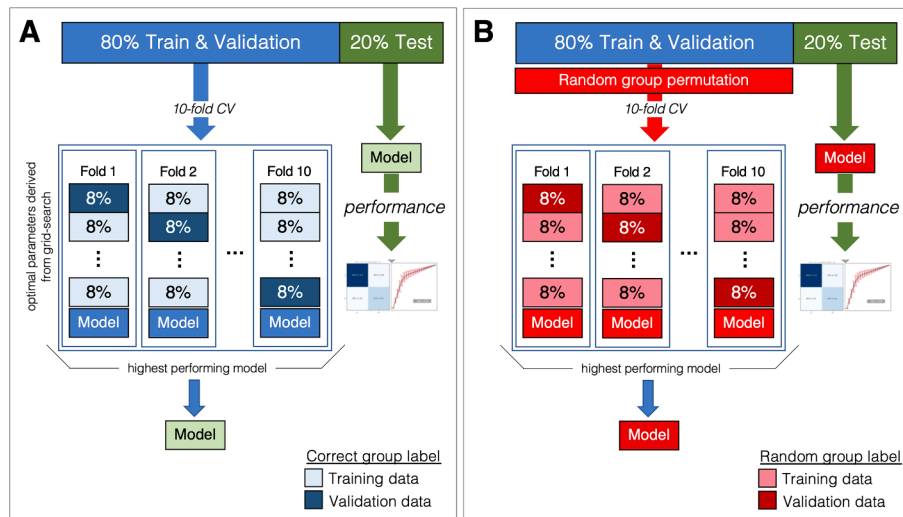
2.7. Pipeline performance evaluation

The performance of our pipeline was evaluated using guidelines recommended by Poldrack et al. (2019). In particular, a 10-fold cross-validation procedure that incorporated a grid search technique was used to evaluate pipeline classification performance (Fig. 1A). More specifically, given an M × N ComBat-harmonized and imbalance corrected *ROI data matrix,* where M (row) is the total number of participants and N (column) is the number of ROI variables, and an M × 1 *group label matrix* that holds the corresponding group label (e.g., HC = 0 and TLE-HS-L = 1) for each participant, the following steps were sequentially applied:

1) The rows of the data and group label matrices were randomly shuffled together (i.e., subject & group label maintained). A 80/20 percent stratified (based on group label) partition split was then

---

Fig. 1. Pipeline performance evaluation. (**A**) Schematic illustration of a 10-fold grid search process to create an optimally trained pipeline with the highest classification accuracy; and (**B**) Schematic illustration of a 10-fold cross-validation process to create a pipeline using shuffled labels to yield a random distribution in order to assess statistical significance between models trained on real vs. permuted (random) data.

applied to the matrices, where 80% were assigned to training and validation data, and 20% became test data. It is important to note that test data were not used to learn the ROI variable scaling parameters or the optimal group classification model parameters (found by the grid search outlined below).

2) Using only the training and validation data split, a 10-fold stratified grid search procedure was applied: one fold was selected as validation data and the remaining nine folds became the training data. An exhaustive grid search (i.e., iterations of all possible model parameter combinations to yield the most reliable set) was performed to estimate the optimal classification model parameters that yielded a pipeline with the highest classification accuracy. The model parameters included in the grid search were: number of nodes in hidden layers one and two, l_2 regularization penalty in hidden regularization layer, and random percent of nodes removed (i.e., dropped) in the hidden dropout layer (**Supplementary Note 1**). This process was repeated until each fold had been selected as the validation fold, resulting in ten trained pipelines (i.e., one for each validation fold).

3) From the ten pipelines created by the grid search procedure above, the optimal pipeline was selected (i.e., one that had the highest classification accuracy) and the optimal model parameters identified (see **Supplementary Notes 1 through 4**).

4) For each subject in the test data matrix, the optimal pipeline was given the participant's set of ROI variables and prompted to predict the participant's group. To assess performance, a 2x2 confusion matrix was constructed using the *predicted* group label and *known* group label (i.e., the real diagnostic group) to calculate: positive predictive value (PPV), negative predictive value (NPV), sensitivity (SEN), specificity (SPC), area under the curve (AUC), and accuracy.

The four steps above were repeated 1,000 times to better estimate the stability of the optimal pipelines (including model parameters) found by the 10-fold cross-validation grid search.

2.8. Assessment of statistical significance

The performance of our pipeline was evaluated using a 10-fold cross-validation procedure that incorporated a *random* group label permutation to estimate the distribution of values under a null hypothesis (Fig. 1**B**). Given an M × N ComBat-harmonized and imbalance corrected *ROI data matrix* and an M × 1 *group label matrix*, the four steps above were carried out, where Steps 1 and 4 were identical to the ones

previously outlined and Steps 2 and 3 differed as follows:

2) The training and validation group label matrix was randomly permuted, and then a 10-fold stratified cross-validation procedure was applied to our pipeline. In this cross-validation, a grid search was not performed. Instead, the highest performing model parameters for the correctly trained pipeline were used.

3) From the ten pipelines created by the cross-validation procedure, the non-optimal pipeline was selected (i.e., one that had the highest classification accuracy with random group label assignments).

As before, these four steps were repeated 1,000 times to better estimate the stability of the non-optimal pipelines found by the 10-fold cross-validation procedure. Using the test data set, the statistical significance of the difference in performance between the optimal and non-optimal pipelines was determined by evaluating how often the mean performance metric of the optimal pipeline was higher than the performance metric of the non-optimal pipeline. For instance, if the average classification accuracy of the optimal pipeline is greater than 98% of the classification accuracies obtained in the non-optimal pipeline, the probability that the optimal pipeline classification accuracy is due to chance is 2% or $p = 0.02$. The same analysis is applied to each performance metric (i.e., accuracy, PPV, NPV, etc.) in our study (Section 3.0).

In order to compare the performance of SVM against DL, we computed the DL accuracy and counted the number of SVM accuracies that were greater than the DL accuracy, then divided by the 1,000 iterations. For between-model comparisons, we report a frequency distribution comparison index (FDCI) where 0.5 demonstrates equipoise in the accuracies of the two approaches (i.e. SVM is as accurate as DL), whereas a FDCI value closer to 1 means that SV outperforms DL on almost every iteration, and vice versa.

2.9. ROI variable selection

For the SVC model, support vector coefficients were used to select the ROI variables that had the largest influence on group classification accuracy. For the DLC model, the backtracking technique (Hazlett et al., 2017; Munsell et al., 2020) was used to select the ROI variables that had the largest influence on group classification accuracy. Values were averaged across the 1000 iterations of model training.

3. Results

The results are organized by imaging modality (structural and diffusion, separately) and the specific group classification comparison that was performed. For each comparison, a 2 × 2 confusion matrix and a summary table displaying the six performance metrics, are shown as mean ± standard deviation (SD) calculated from the one-thousand 2 × 2 confusion matrices generated by the evaluation process (Section 2.6), and the one-thousand random 2 × 2 confusion matrices generated by our evaluation process that permutates the group labels (Section 2.6). For each performance metric, the statistical significance (p-value) is reported.

3.1. Structural data

The performance of our classification pipelines based on structural data is summarized in Fig. 2 and **Supplementary Table 3** (optimal pipeline parameters available in **Supplementary Note 2**). SVM and DL pipelines demonstrated no significant differences in classification accuracy for TLE vs. HC comparisons. There was a 73 ± 2.9% (DL) to 75 ± 3.4% (SVM) accuracy to correctly distinguish all patients with TLE from HC based on GM volumes/thickness, although the accuracy was slightly lower (65–67%) when the model attempted to classify TLE patients based on the specific side of presumed seizure onset (i.e., TLE-HS-L and TLE-HS-R, separately) from HC. While the performance of the latter model was higher than that of a pipeline trained on randomized data, the sensitivity and specificity, especially of DL models, were suboptimal. When classifying TLE patients based on their side of pathology (i.e., TLE-HS-L vs. TLE-HS-R), DL underperformed relative to SVM (77% ± 6.1% vs. 83% ± 3.6%, $p = 0.97$) likely driven by DL's poorer sensitivity/

specificity. As expected, ipsilateral hippocampal volumes had the highest influence on the prediction models in all cases. Other ROIs that showed consistently high importance were the contralateral amygdala, ipsilateral thalamus, and different portions of the bilateral lateral temporal regions and frontal regions.

3.2. Diffusion data

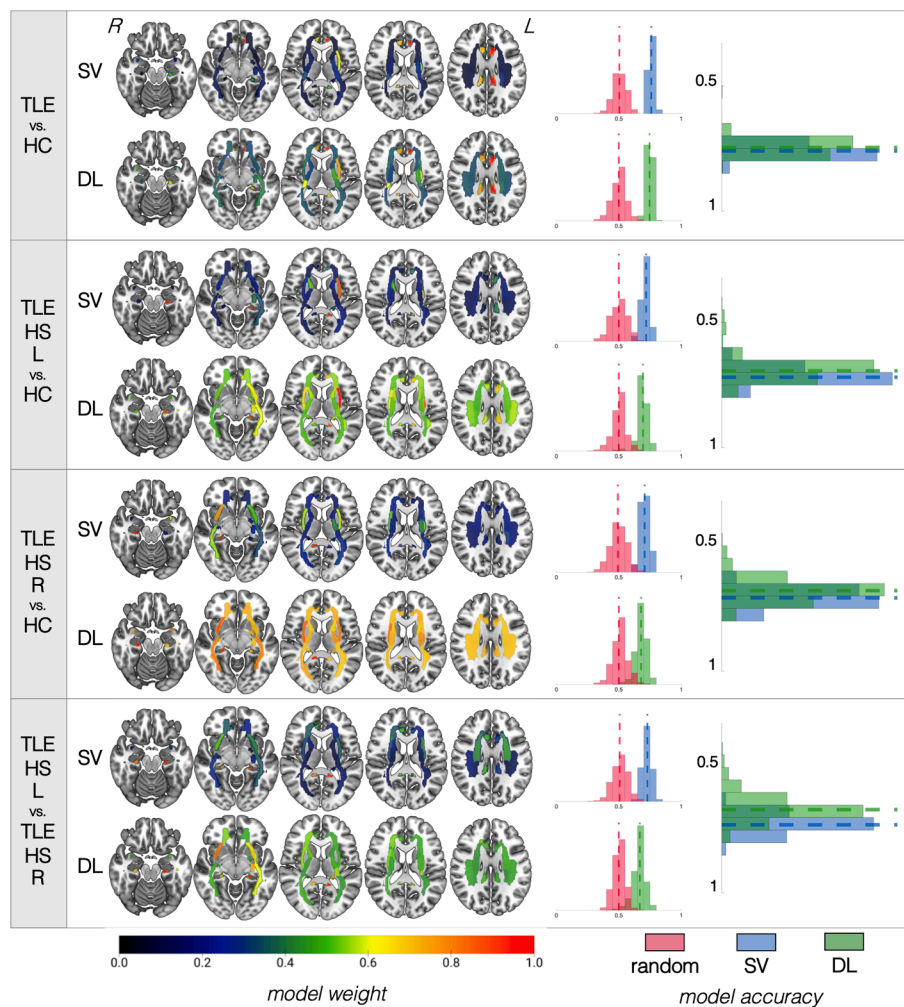### 3.2.1. Fractional anisotropy

The performance of our classification pipelines based on FA data are summarized in Fig. 3 for patients with hippocampal sclerosis and in Fig. 4 for patients without hippocampal sclerosis, as well as **Supplementary Table 4** (for optimal pipeline parameters, please see **Supplementary Note 3**).

Patients with hippocampal sclerosis: In all comparisons of TLE groups vs. HC, the SVM and DL pipelines demonstrated no detectable differences in classification accuracy. There was a 74 ± 2.7% (DL) to 76 ± 5.4% (SVM) accuracy to correctly distinguish all patients with TLE-HS from HC based on FA values. Similar accuracies (71–72%) were found when the models attempted to classify TLE patients based on the specific side of presumed seizure onset (i.e., TLE-HS-L and TLE-HS-R, separately) from HC. This model showed significantly higher sensitivity/specificity than a classifier trained on randomized data for SVM but not for DL. Lateralization of TLE-HS patients demonstrated 73% ± 3.4% accuracy with SVM and 66% ± 5.2% with DL, demonstrating the relative suboptimal performance of the latter (p = 0.97), again driven by the lower sensitivity/specificity. In all cases, the bilateral (ipsilateral > contralateral) parahippocampal cingulate bundles had the most weight. The external capsule appeared to have more prominent weight for classification of TLE groups vs. HC than for lateralization.
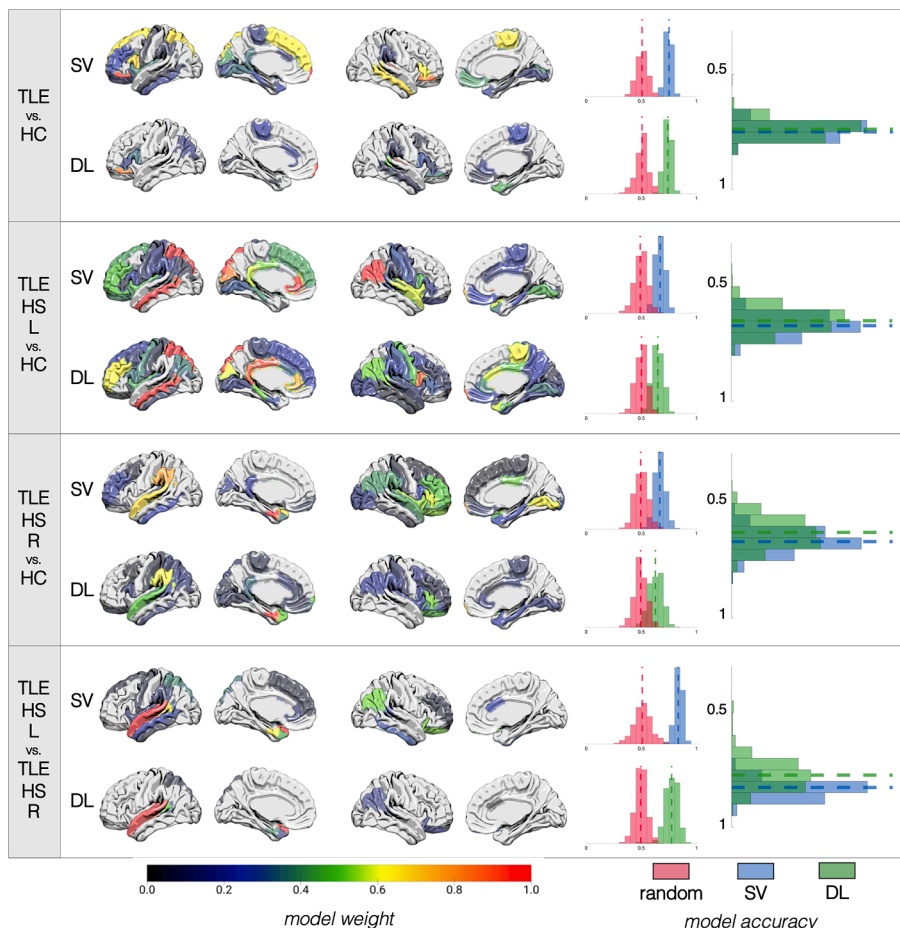
Fig. 2. Summary of results for models based on structural data. For each of the four classifications (identified in the leftmost column), the top row shows the results for the support vector (SV) approach and the bottom row shows the results for the deep learning (DL) approach. Cortex projections are shown, from left to right, overlaid on a left lateral, left medial, right lateral, and right medial view, respectively. Projections correspond to model weights for each region of interest (ROI) based on the Desikan-Killiany atlas colored based on the colormap at the bottom of the figure and normalized from 0 to 1, where higher values (more red regions) correspond to the most influential ROIs for that particular model. For each classification approach, a distribution of accuracies is shown to the right for permutated labels (red) or real labels (blue for SV, green for DL). The rightmost graph compares the accuracy of SV vs. DL against each other. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 3. Summary of results for models based on fractional anisotropy data for patients with hippocampal sclerosis. For each of the four classifications (identified in the leftmost column), the top row shows the results for the support vector (SV) approach and the bottom row shows the results for the deep learning (DL) approach. White matter tracts are shown on axial slices from ventral to dorsal. The orientation corresponds to radiological reference (i.e., the left side of the axial slice corresponds to the right side of the brain, and vice versa). Each bundle corresponds to a region of interest (ROI) based on the Johns Hopkins University atlas colored based on the colormap at the bottom of the figure and normalized from 0 to 1, where higher values (more red regions) correspond to the most influential ROIs for that particular model. For each classification approach, a distribution of accuracies is shown to the right for permutated labels (red) or real labels (blue for SV, green for DL). The rightmost graph compares the accuracy of SV vs. DL against each other. Notice that models for which ROI weights have less variable distribution (i.e., all ROIs are relatively of equal importance for classification, as shown by a similar color range across all regions) tend to have worse performance accuracies. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Patients without hippocampal sclerosis: In all cases, SVM and DL pipelines demonstrated no significant differences in classification accuracy ($p = 0.44$ to 0.83). There was a $60 \pm 4\%$ (DL) to $62\% \pm 4\%$ (SVM) accuracy to correctly distinguish all patients with TLE-NS from HC based on FA values. The performance was lower (53–63%) when the models attempted to classify TLE patients based on the specific side of presumed seizure onset (i.e., TLE-NS-L and TLE-NS-R, separately) from HC, and these models were not significantly more accurate than models trained on randomized data. Similarly, attempts to predict lateralization of TLE-NS patients demonstrated low accuracy (51% DL, 56% SVM) which did not outperform models trained on randomized data. The bilateral cingulate bundles showed the most prominence for SVM and DL models classifying all TLE-NS independently of side vs. HC, but these tracts were not as important for other models. The ipsilateral sagittal stratum was important in classifying TLE-NS patients from each side against controls. For TLE-NS-R vs. HC in particular, the ipsilateral cingulate, contralateral fornix, and ipsilateral inferior fronto-occipital fasciculus, as well as the splenium of the corpus callosum, all appeared to influence the SVM model more heavily than for other comparisons.
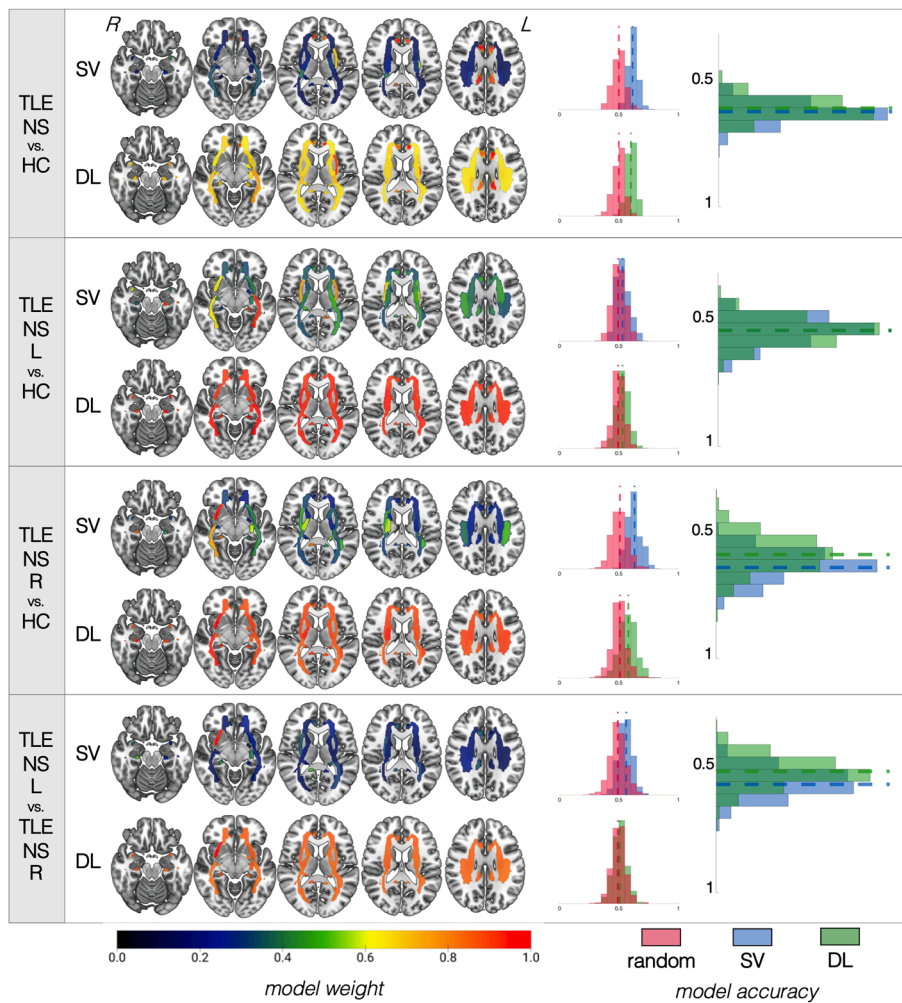
3.2.2. Radial diffusivity

The performance of our classification pipelines based on RD are summarized in Fig. 5 for patients with hippocampal sclerosis and Fig. 6 for patients without hippocampal sclerosis, as well as **Supplementary Table 5** (for optimal pipeline parameters, please see **Supplementary Note 4**).

Patients with hippocampal sclerosis: In all TLE vs. HC

comparisons, SVM and DL pipelines demonstrated no significant differences in classification accuracy. There was a $71 \pm 5.9\%$ (DL) to $73 \pm 2.5\%$ (SVM) accuracy for identifying all patients with TLE-HS versus HC based on RD values. Slightly lower accuracies (67–69%) were found when the models attempted to classify TLE patients based on the specific side of presumed seizure onset (i.e., TLE-HS-L and TLE-HS-R, separately) from HC. These models showed significantly higher sensitivity/specificity than a classifier trained on randomized data for SVM but not for DL. Lateralization of TLE-HS patients demonstrated $68\% \pm 4.1\%$ accuracy with SVM but only $55\% \pm 5.7\%$ with DL, which had relatively suboptimal performance ($p = 1.0$), again driven by the lower sensitivity/specificity of the latter. In all cases, the bilateral (ipsilateral > contralateral) cingulate bundles had the most weight. The external capsule appeared to have a more prominent weight for classification of TLE groups vs. HC than for lateralization. For lateralization in particular, the superior corona radiata bilaterally demonstrated higher weights than for TLE vs. HC comparisons. Bilateral inferior fronto-occipital fasciculi were particularly influential for the SVM model of TLE-HS-R vs. HC classification.

Patients without hippocampal sclerosis: In all cases, SVM and DL pipelines demonstrated no significant differences in classification accuracy. There was a $59 \pm 4.5\%$ (DL) to $62\% \pm 4\%$ (SVM) accuracy to correctly identify all patients with TLE-NS from HC based on RD values. The performance was lower (54–63%) when the models attempted to classify TLE patients based on the specific side of presumed seizure onset (i.e., TLE-HS-L and TLE-HS-R, separately) from HC, which were not significantly more accurate than models trained on randomized data.

Fig. 4. Summary of results for models based on fractional anisotropy data for patients without hippocampal sclerosis. For each of the four classifications (identified in the leftmost column), the top row shows the results for the support vector (SV) approach and the bottom row shows the results for the deep learning (DL) approach. White matter tracts are shown on axial slices from ventral to dorsal. The orientation corresponds to radiological reference (i.e., the left side of the axial slice corresponds to the right side of the brain, and vice versa). Each bundle corresponds to a region of interest (ROI) based on the Johns Hopkins University atlas colored based on the colormap at the bottom of the figure and normalized from 0 to 1, where higher values (more red regions) correspond to the most influential ROIs for that particular model. For each classification approach, a distribution of accuracies is shown to the right for permutated labels (red) or real labels (blue for SV, green for DL). The rightmost graph compares the accuracy of SV vs. DL against each other. Notice that models for which ROI weights have less variable distribution (i.e., all ROIs are relatively of equal importance for classification, as shown by a similar color range across all regions) tend to have worse performance accuracies. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Similarly, attempts to classify the lateralization of TLE-NS patients demonstrated low accuracy (55% DL, 56% SV) which was not better than chance. Bilateral cingulate bundles showed the most prominence for SVM and DL models classifying all TLE-NS independently of side vs. HC, but these tracts were not as important for other models. The ipsilateral sagittal stratum was important in classifying TLE-NS patients from each side against controls. Bilateral (ipsilateral > contralateral) cingulum and external capsule bundles contributed heavily to all classification models. The inferior fronto-occipital fasciculus was more important particularly for TLE-NS-R both vs. TLE-NS-L and also vs. HC. The sagittal stratum was influential for comparisons against HC ipsilaterally, and bilaterally for lateralization.

4. Discussion

This study leveraged the large dataset from the ENIGMA-Epilepsy consortium to evaluate the performance of conventional machine learning (SVM) and DL in distinguishing patients with TLE versus healthy controls, as well as lateralizing TLE using individual structural and diffusion data at an ROI level. The overarching goal of the study was to demonstrate the feasibility of applying machine learning and artificial intelligence frameworks to large cohorts of clinically-obtained imaging data. We observed that 1) SVM was more consistently accurate than DL in outperforming random models, 2) models to discriminate TLE patients from healthy controls performed better than models to lateralize the seizure focus in TLE patients, 3) structural and diffusion-based models showed overall similar classification accuracies, 4)
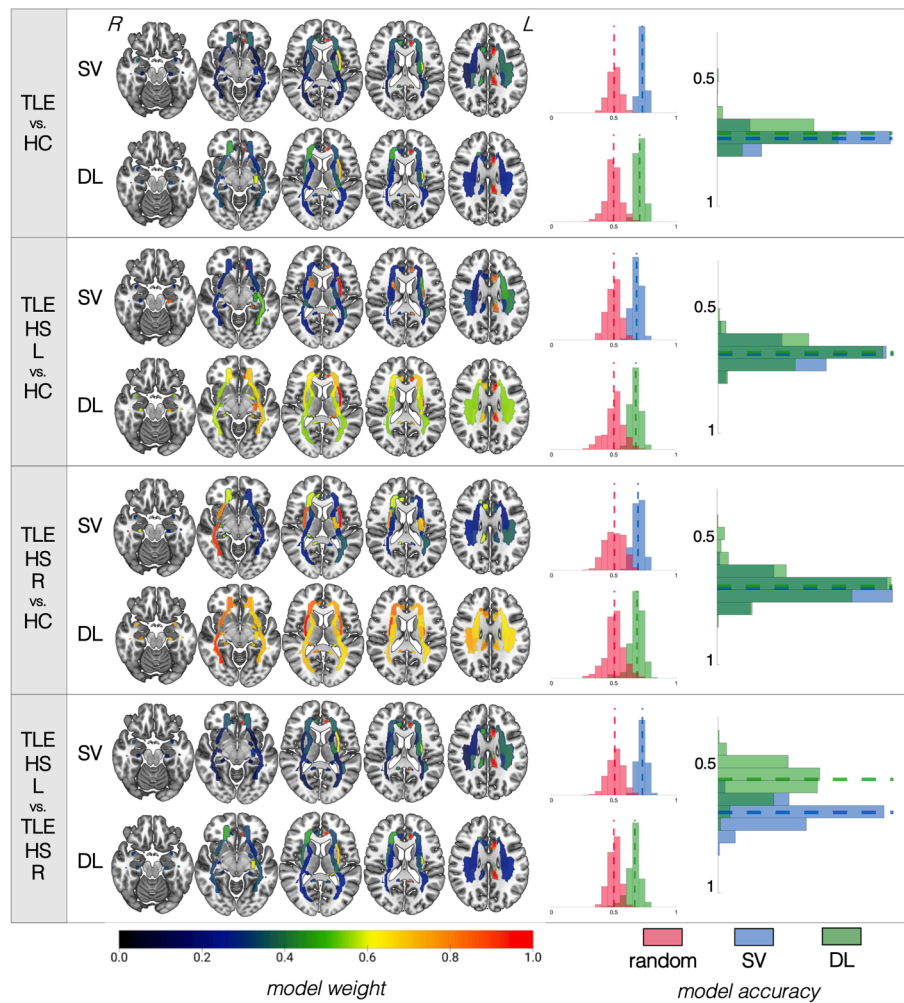
classification models of patients with HS were more accurate than models aimed at stratifying non-lesional patients, and 5) many ipsilateral and contralateral white matter tracts contributed to model classifications.

4.1. ROI level data

While this is the largest artificial intelligence study of quantitative imaging in epilepsy, our study was based on ROI-level data. Specifically, for each subject, there were 152 data points for structural imaging (derived from 84 ROIs) and 41 data points for diffusion imaging, all representing the average values within each of the ROIs. This is a much lower resolution compared to voxel-wise data, in which each subject would have millions of data points per imaging modality. The motivations for conducting an ROI-level artificial intelligence study are: 1) the available ROI-level ENIGMA dataset yields a substantially larger sample size than single or even multi-site cohorts; 2) the identification of promising (statistically significant, i.e., better than chance) results with ROI-level data motivates future studies with higher-resolution data; and 3) we can find out if ROI-level analysis of large sample sizes is sufficient for near-perfect classification; in other words, it can help evaluate if the determinant for classification accuracy in machine learning models is merely sample size.

We observed promising classification accuracies for many imaging modalities and metrics (detailed below). They were, however, not perfect. As such, it is possible to derive two initial conclusions: 1) artificial intelligence has a growing potential in the field of epilepsy, and 2)

Fig. 5. Summary of results for models based on radial diffusivity data for patients with hippocampal sclerosis. For each of the four classifications (identified in the leftmost column), the top row shows the results for the support vector (SV) approach and the bottom row shows the results for the deep learning (DL) approach. White matter tracts are shown on axial slices from ventral to dorsal. The orientation corresponds to radiological reference (i.e., the left side of the axial slice corresponds to the right side of the brain, and vice versa). Each bundle corresponds to a region of interest (ROI) based on the Johns Hopkins University atlas colored based on the colormap at the bottom of the figure and normalized from 0 to 1, where higher values (more red regions) correspond to the most influential ROIs for that particular model. For each classification approach, a distribution of accuracies is shown to the right for permutated labels (red) or real labels (blue for SV, green for DL). The rightmost graph compares the accuracy of SV vs. DL against each other. Notice that models for which ROI weights have less variable distribution (i.e., all ROIs are relatively of equal importance for classification, as shown by a similar color range across all regions) tend to have worse performance accuracies. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

further refinements will likely not depend on larger sample sizes of ROI-level data, but on higher resolution imaging, including voxel- and vertex-level data.

4.2. SVM vs. DL model performance

SVM and DL showed similar accuracy for most models stratifying patients from healthy controls. However, several SVM-based models performed better in terms of sensitivity and specificity relative to models trained on randomized data, whereas DL-based models for the same imaging modalities showed non-superior performance. In addition, there was overall relative suboptimal DL performance relative to SVM particularly for lateralization of TLE groups across imaging modalities. SVM models may be moderately reliable for classification and lateralization of TLE based on structural and diffusion sequences, as previously shown by other smaller-scale studies using modalities such as resting state functional MRI (Zhang et al., 2012).
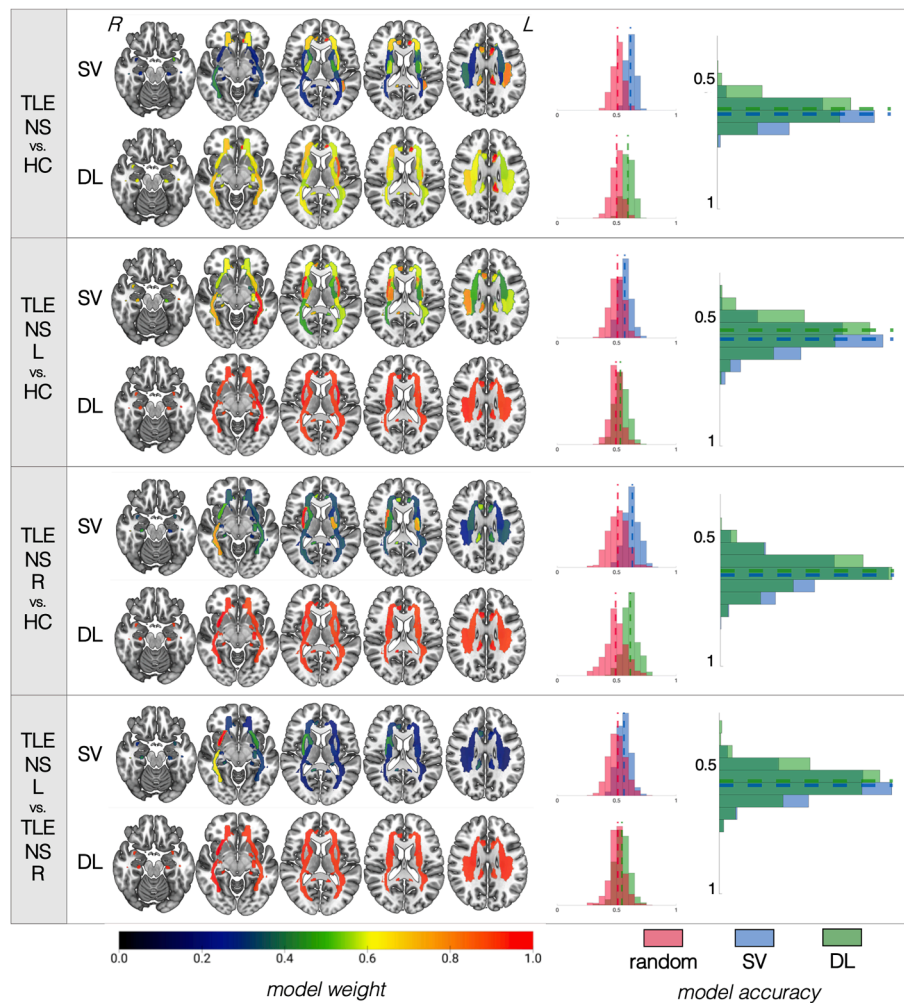
One reason for DL's lower performance might be related to the type of features used in our analyses. The computation of whole-ROI averages from multiple voxels can be regarded as a coarse approach with information inherently being lost. This approach may be suboptimal for DL techniques as these can make use of fine-grained features such as edges or clusters within a complete dataset that has not been aggregated into a fixed set of ROIs. In such cases, the performance of a simple linear decision boundary found by a support vector classification technique is likely to be just as accurate, or possibly more accurate, than a complex decision network found by a deep learning approach. These findings highlight the need to enter different types of features with varying degrees of resolution (i.e. fine-grained for DL and course-grained for SV).

4.3. TLE diagnosis vs. Lateralization

Our machine learning models were similar in distinguishing TLE patients from healthy controls and were less reliable in lateralizing patients based on their side of TLE, particularly for SVM. Consistently across SVM and DL, however, when classifying non-lesional patients, the models performed slightly worse for lateralization than diagnostic stratification. Structural-based models had a superior accuracy for detection over lateralization of TLE, particularly for SVM (83% detection vs. 67% lateralization, on average). The issue of epilepsy lateralization was one of the earliest to be addressed by machine learning approaches with modalities including positron emission tomography (PET) (Lee et al., 2000), photon emission computed tomography (SPECT) (Lopes et al., 2010), and diffusion magnetic resonance techniques (Davoodi-Bojd et al., 2016) including diffusion tensor (Ahmadi et al., 2009; Concha et al., 2012; Kamiya et al., 2016) and kurtosis (Del Gaizo et al., 2017) imaging, with mean accuracies overall similar to those found in this study. Future studies might enhance lateralization accuracy by building classifiers based on multi-modal features, particularly for non-lesional patients, whose lateralization was suboptimal and often not better than that of a random model.

Fig. 6. Summary of results for models based on radial diffusivity data for patients without hippocampal sclerosis. For each of the four classifications (identified in the leftmost column), the top row shows the results for the support vector (SV) approach and the bottom row shows the results for the deep learning (DL) approach. White matter tracts are shown on axial slices from ventral to dorsal. The orientation corresponds to radiological reference (i.e., the left side of the axial slice corresponds to the right side of the brain, and vice versa). Each bundle corresponds to a region of interest (ROI) based on the Johns Hopkins University atlas colored based on the colormap at the bottom of the figure and normalized from 0 to 1, where higher values (more red regions) correspond to the most influential ROIs for that particular model. For each classification approach, a distribution of accuracies is shown to the right for permutated labels (red) or real labels (blue for SV, green for DL). The rightmost graph compares the accuracy of SV vs. DL against each other. Notice that models for which ROI weights have less variable distribution (i.e., all ROIs are relatively of equal importance for classification, as shown by a similar color range across all regions) tend to have worse performance accuracies. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4.4. Structural vs. diffusion models

Models trained with features derived from structural and diffusion imaging performed relatively similarly. One exception was, as stated above, noticeably higher accuracy of structural than diffusion models for lateralization of lesional patients. Both types of modalities provide valuable information for classification, but multimodal approaches may enhance model performance. A combination of data types could be explored both "within modality" (for example, combining FA, RD, MD, and other diffusion scalar metrics) but also "across modalities" (for example, combining regional volumetrics with FA, etc). In this study, diffusion data obtained at the ROI-level condensed large amounts of white matter bundle information into regional data points. Utilizing tractography metrics that do not conform to predefined white matter bundles (i.e., ROI-to-ROI connectivity) might also yield higher accuracies by providing higher resolution at the whole-brain level, unmasking patterns of abnormal network organization undetected by nodal measures. Similarly, future studies could combine features derived from graph theory metrics or other connectome-based approaches, which could serve as topological biomarkers for the detection and lateralization of TLE.

4.5. Lesional vs. non-lesional patients

As expected, models performed better in lesional than non-lesional patients. The expectation of machine learning models with so-called 'MRI-negative' patients is that these computational techniques will be

able to detect features otherwise visually ignored or undetectable by human experts to enhance classification accuracy. We did not have access to non-lesional patients for structural data, so our interpretation must be taken with caution. However, based on the diffusion models, it is evident that regional information may be suboptimal to detect and lateralize non-lesional patients, with accuracies mostly in the 53–63% range. This is in line with previous studies in focal cortical dysplasia showing similar accuracy (58%) in detecting focal cortical dysplasia lesions in MRI studies deemed to be 'negative' by human experts (Ahmed et al., 2015; Hong et al., 2014). Importantly, we must acknowledge that there is a bias inherent to the proportion of lesional vs. non-lesional patients in adults with TLE, which invariably features a larger sample of the former group. A larger number of patients has the potential to allow for more fine-grained modeling based on more samples for training. Future studies could address this issue by testing the accuracies of samples of similar size in each group.

4.6. Model features

Across different models, the features with maximal importance for classification accuracy included structures in the limbic region (e.g., cingulate bundles and sagittal stratum) as well as extra-limbic portions. In structural models, ipsilateral hippocampal volumes had the highest weight with variable influence of other medial temporal structures, such as the amygdala, as well as lateral temporal regions and extra temporal areas (including the thalamus), all of which have been shown to have volume abnormalities using *meta*-analytical group comparisons (Whelan

et al., 2018). Similarly, diffusion models demonstrated that the most influential white matter ROIs were along limbic (e.g., cingulate bundles, sagittal stratum) and extra-limbic (e.g., external capsule) sub-networks, aligned with recent mega-analysis across epilepsy syndromes by the ENIGMA-Epilepsy consortium (Hatton et al., 2020). These findings highlight the now well-established observation that plastic changes in focal epilepsy occur beyond the presumed area of seizure focus, stressing the nature of epilepsy as a disorder characterized by aberrant brain network reorganization (Engel et al., 2013; Gleichgerrcht and Bonilha, 2017; Richardson, 2012; Tavakol et al., 2019).

Limitations

The initial cohort of ENIGMA-Epilepsy patients included ROI-based data only. Because post-processing was performed at each site following structured guidelines and quality checks from the consortium protocols, the data collected were exclusively at the ROI level. We believe that the suboptimal performance of some of the classification models reported above is being driven by the input data existing at the ROI level, which summarizes potentially heterogeneous information into a single value. The trade-off to this limitation, however, was the large volume of patients with TLE from epilepsy centers across the world. One alternative approach in the absence of voxel-level data would be to implement finer resolution atlases that yield smaller ROIs, hence counting on a more rich dataset per patient where each individual value reflects a more constrained anatomical region. As such, this large multi-site, multi-nationality representation of TLE provides the most comprehensive representation of the disease to date. Nonetheless, such a large database cohort carries the potential to be affected by confounders intrinsic to the retrospective nature of our design, including variations in type of scanner, variability in local quality assurance protocols, acquisition protocols, and regional clinical diagnostic pipelines.

We addressed imaging heterogeneity by applying a well-established data harmonization technique validated in diffusion tensor imaging data (Fortin et al., 2017) and applied it to multi-site cohorts of epilepsy patients (Hatton et al., 2020). By decreasing imaging data heterogeneity, we believe that the results from this study may be more generalizable. However, there are a few limitations to this approach. First, because ComBat is applied to the entire data set, and then split into a train, validation, and test dataset, a bias is likely introduced into the classification pipeline. More specifically, because the harmonization parameters learned by ComBat included the test data, the test data is not completely unseen which may slightly enhance classification performance. Second, ROI data from the same site may appear in both the training and test data set, so a site bias may be introduced into the classification. However, since ComBat was applied to the entire dataset prior to pipeline construction, multi-site differences were minimized (as best as possible) and classification performance enhancements due to site bias were likely mitigated. More specifically, relatively lower classification accuracy (i.e. 50 to 70%) could be an indicator of site bias removal.

In addition, it is important to consider other non-imaging related sources that may further limit model accuracy. For instance, the overall underperformance of classification models for lateralization of TLE may be negatively influenced by how left vs. right disease is determined (i.e., based on EEG and/or imaging findings). However, future studies could use a different approach by comparing accuracies exclusively among those who achieved seizure freedom (i.e., Engel class I or ILAE 1 outcome), a proxy for "gold standard" confirmation of laterality (for further discussion, see Liu et al., 2016).

The limitations of this study also highlight the importance of multimodal approaches. Because data collection for ENIGMA-Epilepsy occurred in different waves for structural and diffusion images, we were not able to cross-reference anonymized data across cohorts in order to test bimodal inputs to train the classifier. Efforts are currently ongoing for each consortium to identify the overlap between structural and diffusion cohorts to enable multimodal testing. In addition, ENIGMA-Epilepsy also includes patients with generalized genetic epilepsy and extratemporal focal epilepsy. Similar studies with these populations may highlight classifier accuracy drivers common for all epilepsy syndromes and also distinct patterns that inform on syndrome-specific characteristics.

5. Conclusions

Artificial intelligence might allow for the classification and lateralizing of TLE using ROI-level data, with moderate accuracy. Cortical thickness and surface area were equivalent in accuracy compared with fractional anisotropy and radial diffusivity, and classification was better among patients with hippocampal atrophy. DL was not superior to SVM at ROI-level data. Extra-hippocampal features were important in all classification models, suggesting that machine learning can use information beyond regions that are visually identified in human qualitative analyses.

These results suggest that machine learning approaches can aid in the classification of TLE (which could potentially aid in radiological diagnosis). This process may be improved by employing higher resolution imaging data (i.e., pre-processed images with no smoothing or manipulations) and multimodal approaches. This study will be a driver for ENIGMA-Epilepsy and other large data consortia to centralize and harmonize raw data for machine learning analyses.

6. Funding and acknowledgements

Credit Author Statement

Data curation: N.Ba., A.B., N.Be., L.B., E.G., F.C., C.De., O.D., M.D., N.K.F., A.G., R.G., R.Kä., S.S.K., P.K., A.L., S.L., C.R.M., S.M., T.J., H.R. P., M.P.R., T.R., S.M.S., H.S.-Z., P.S., B.W., R.W., Y.Y., J.Z.

Formal analysis: E.G., L.B., S.N.H., P.K., C.R.M., B.C.M.

Methodology: E.G., B.C.M., L.B., S.A., B.Ben., M.E.C., M.D., S.N.H., S. S.K., R.Ko., B.A.K.K., M.L., P.M., C.R.M., H.R.P., J.C.P., R.R.-C., B.S., P. N.T., A.E.V., L.V., C.L.Y.

Writing – original draft: L.B., E.G., C.R.M., B.C.M.

Writing – reviewing & editing: S.A., A.A., B.Ben., A.B., N.Be., B.Ber., L.B., M.E.C., F.C., L.C., M.D., J.S.D., N.K.F., M.G., A.G., E.G., B.G., S.S.K., P.K., B.A.K.K., S.L., S.La., M.L., E.L., M.M., C.R.M., B.C.M., T.J., H.R.P., M.P.R., S.M.S., D.J.S., P.S., S.I.T., P.M.T., L.V., F.v.P., S.B.V.

Project administration: C.R.M., S.M.S.

Declaration of Competing Interest

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.nicl.2021.102765.

References

Abbasi, B., Goldenholz, D.M., 2019. Machine learning applications in epilepsy. Epilepsia 60, 2037–2047.

Adams, H.H., Hibar, D.P., Chouraki, V., Stein, J.L., Nyquist, P.A., Renteria, M.E., Trompet, S., Arias-Vasquez, A., Seshadri, S., Desrivieres, S., Beecham, A.H., Jahanshad, N., Wittfeld, K., Van der Lee, S.J., Abramovic, L., Alhusaini, S., Amin, N., Andersson, M., Arfanakis, K., Aribisala, B.S., Armstrong, N.J., Athanasiu, L., Axelsson, T., Beiser, A., Bernard, M., Bis, J.C., Blanken, L.M., Blanton, S.H., Bohlken, M.M., Boks, M.P., Bralten, J., Brickman, A.M., Carmichael, O., Chakravarty, M.M., Chauhan, G., Chen, Q., Ching, C.R., Cuellar-Partida, G., Braber, A.D., Doan, N.T., Ehrlich, S., Filippi, I., Ge, T., Giddaluru, S., Goldman, A.L., Gottesman, R.F., Greven, C.U., Grimm, O., Griswold, M.E., Guadalupe, T., Hass, J., Haukvik, U.K., Hilal, S., Hofer, E., Hoehn, D., Holmes, A.J., Hoogman, M., Janowitz, D., Jia, T., Kasperaviciute, D., Kim, S., Klein, M., Kraemer, B., Lee, P.H., Liao, J., Liewald, D.C., Lopez, L.M., Luciano, M., Macare, C., Marquand, A., Matarin, M., Mather, K.A., Mattheisen, M., Mazoyer, B., McKay, D.R., McWhirter, R., Milaneschi, Y., Mirza-Schreiber, N., Muetzel, R.L., Maniega, S.M., Nho, K., Nugent, A.C., Loohuis, L.M., Oosterlaan, J., Papmeyer, M., Pappa, I., Pirpamer, L., Pudas, S., Putz, B., Rajan, K.B., Ramasamy, A., Richards, J.S., Risacher, S.L., Roiz-Santianez, R., Rommelse, N., Rose, E.J., Royle, N.A., Rundek, T., Samann, P.G., Satizabal, C.L., Schmaal, L., Schork, A.J., Shen, L., Shin, J., Shumskaya, E., Smith, A. V., Sprooten, E., Strike, L.T., Teumer, A., Thomson, R., Tordesillas-Gutierrez, D., Toro, R., Trabzuni, D., Vaidya, D., Van der Grond, J., Van der Meer, D., Van Donkelaar, M.M., Van Eijk, K.R., Van Erp, T.G., Van Rooij, D., Walton, E., Westlye, L.

T., Whelan, C.D., Windham, B.G., Winkler, A.M., Woldehawariat, G., Wolf, C., Wolfers, T., Xu, B., Yanek, L.R., Yang, J., Zijdenbos, A., Zwiers, M.P., Agartz, I., Aggarwal, N.T., Almasy, L., Ames, D., Amouyel, P., Andreassen, O.A., Arepalli, S., Assareh, A.A., Barral, S., Bastin, M.E., Becker, D.M., Becker, J.T., Bennett, D.A., Blangero, J., van Bokhoven, H., Boomsma, D.I., Brodaty, H., Brouwer, R.M., Brunner, H.G., Buckner, R.L., Buitelaar, J.K., Bulayeva, K.B., Cahn, W., Calhoun, V. D., Cannon, D.M., Cavalleri, G.L., Chen, C., Cheng, C.Y., Cichon, S., Cookson, M.R., Corvin, A., Crespo-Facorro, B., Curran, J.E., Czisch, M., Dale, A.M., Davies, G.E., De Geus, E.J., De Jager, P.L., de Zubicaray, G.I., Delanty, N., Depondt, C., DeStefano, A. L., Dillman, A., Djurovic, S., Donohoe, G., Drevets, W.C., Duggirala, R., Dyer, T.D., Erk, S., Espeseth, T., Evans, D.A., Fedko, I.O., Fernandez, G., Ferrucci, L., Fisher, S.E., Fleischman, D.A., Ford, I., Foroud, T.M., Fox, P.T., Francks, C., Fukunaga, M., Gibbs, J.R., Glahn, D.C., Gollub, R.L., Goring, H.H., Grabe, H.J., Green, R.C., Gruber, O., Gudnason, V., Guelfi, S., Hansell, N.K., Hardy, J., Hartman, C.A., Hashimoto, R., Hegenscheid, K., Heinz, A., Le Hellard, S., Hernandez, D.G., Heslenfeld, D.J., Ho, B.C., Hoekstra, P.J., Hoffmann, W., Hofman, A., Holsboer, F., Homuth, G., Hosten, N., Hottenga, J.J., Hulshoff Pol, H.E., Ikeda, M., Ikram, M.K., Jack Jr., C.R., Jenkinson, M., Johnson, R., Jonsson, E.G., Jukema, J.W., Kahn, R.S., Kanai, R., Kloszewska, I., Knopman, D.S., Kochunov, P., Kwok, J.B., Lawrie, S.M., Lemaitre, H., Liu, X., Longo, D.L., Longstreth Jr., W.T., Lopez, O.L., Lovestone, S., Martinez, O., Martinot, J.L., Mattay, V.S., McDonald, C., McIntosh, A.M., McMahon, K.L., McMahon, F.J., Mecocci, P., Melle, I., Meyer-Lindenberg, A., Mohnke, S., Montgomery, G.W., Morris, D.W., Mosley, T.H., Muhleisen, T.W., Muller-Myhsok, B., Nalls, M.A., Nauck, M., Nichols, T.E., Niessen, W.J., Nothen, M. M., Nyberg, L., Ohi, K., Olvera, R.L., Ophoff, R.A., Pandolfo, M., Paus, T., Pausova, Z., Penninx, B.W., Pike, G.B., Potkin, S.G., Psaty, B.M., Reppermund, S., Rietschel, M., Roffman, J.L., Romanczuk-Seiferth, N., Rotter, J.I., Ryten, M., Sacco, R.L., Sachdev, P.S., Saykin, A.J., Schmidt, R., Schofield, P.R., Sigurdsson, S., Simmons, A., Singleton, A., Sisodiya, S.M., Smith, C., Smoller, J.W., Soininen, H., Srikanth, V., Steen, V.M., Stott, D.J., Sussmann, J.E., Thalamuthu, A., Tiemeier, H., Toga, A.W., Traynor, B.J., Troncoso, J., Turner, J.A., Tzourio, C., Uitterlinden, A.G., Hernandez, M.C., Van der Brug, M., Van der Lugt, A., Van der Wee, N.J., Van Duijn, C.M., Van Haren, N.E., Van, T.E.D., Van Tol, M.J., Vardarajan, B.N., Veltman, D.J., Vernooij, M.W., Volzke, H., Walter, H., Wardlaw, J.M., Wassink, T.H., Weale, M.E., Weinberger, D.R., Weiner, M.W., Wen, W., Westman, E., White, T., Wong, T.Y., Wright, C.B., Zielke, H.R., Zonderman, A.B., Deary, I.J., DeCarli, C., Schmidt, H., Martin, N.G., De Craen, A.J., Wright, M.J., Launer, L.J., Schumann, G., Fornage, M., Franke, B., Debette, S., Medland, S.E., Ikram, M.A., Thompson, P.M., 2016. Novel genetic loci underlying human intracranial volume identified through genome-wide association. Nat. Neurosci. 19, 1569–1582.

Ahmadi, M.E., Hagler Jr., D.J., McDonald, C.R., Tecoma, E.S., Iragui, V.J., Dale, A.M., Halgren, E., 2009. Side matters: diffusion tensor imaging tractography in left and right temporal lobe epilepsy. AJNR Am. J. Neuroradiol. 30, 1740–1747.

Ahmed, B., Brodley, C.E., Blackmon, K.E., Kuzniecky, R., Barash, G., Carlson, C., Quinn, B.T., Doyle, W., French, J., Devinsky, O., Thesen, T., 2015. Cortical feature analysis and machine learning improves detection of "MRI-negative" focal cortical dysplasia. Epilepsy Behav. 48, 21–28.

Babb, T., Brown, W., 1987. Pathological findings in epilepsy. In: Engel, J. (Ed.), Surgical treatment of the epilepsies. Raven, New York, pp. 511–540.

Berg, A.T., Berkovic, S.F., Brodie, M.J., Buchhalter, J., Cross, J.H., van Emde Boas, W., Engel, J., French, J., Glauser, T.A., Mathern, G.W., Moshe, S.L., Nordli, D., Plouin, P., Scheffer, I.E., 2010. Revised terminology and concepts for organization of seizures and epilepsies: report of the ILAE Commission on Classification and Terminology, 2005–2009. Epilepsia 51, 676–685.

Bernhardt, B.C., Bernasconi, N., Concha, L., Bernasconi, A., 2010. Cortical thickness analysis in temporal lobe epilepsy: reproducibility and relation to outcome. Neurology 74, 1776–1784.

Bernhardt, B.C., Hong, S.J., Bernasconi, A., Bernasconi, N., 2015. Magnetic resonance imaging pattern learning in temporal lobe epilepsy: classification and prognostics. Ann. Neurol. 77, 436–446.

Bernhardt, B.C., Worsley, K.J., Besson, P., Concha, L., Lerch, J.P., Evans, A.C., Bernasconi, N., 2008. Mapping limbic network organization in temporal lobe epilepsy using morphometric correlations: insights on the relation between mesiotemporal connectivity and cortical atrophy. Neuroimage 42, 515–524.

Blanc, F., Martinian, L., Liagkouras, I., Catarino, C., Sisodiya, S.M., Thom, M., 2011. Investigation of widespread neocortical pathology associated with hippocampal sclerosis in epilepsy: a postmortem study. Epilepsia 52, 10–21.

Blumcke, I., Spreafico, R., Haaker, G., Coras, R., Kobow, K., Bien, C.G., Pfafflin, M., Elger, C., Widman, G., Schramm, J., Becker, A., Braun, K.P., Leijten, F., Baayen, J.C., Aronica, E., Chassoux, F., Hamer, H., Stefan, H., Rossler, K., Thom, M., Walker, M.C., Sisodiya, S.M., Duncan, J.S., McEvoy, A.W., Pieper, T., Holthausen, H., Kudernatsch, M., Meencke, H.J., Kahane, P., Schulze-Bonhage, A., Zentner, J., Heiland, D.H., Urbach, H., Steinhoff, B.J., Bast, T., Tassi, L., Lo Russo, G., Ozkara, C., Oz, B., Krsek, P., Vogelgesang, S., Runge, U., Lerche, H., Weber, Y., Honavar, M., Pimentel, J., Arzimanoglou, A., Ulate-Campos, A., Noachtar, S., Hartl, E., Schijns, O., Guerrini, R., Barba, C., Jacques, T.S., Cross, J.H., Feucht, M., Muhlebner, A., Grunwald, T., Trinka, E., Winkler, P.A., Gil-Nagel, A., Toledano Delgado, R., Mayer, T., Lutz, M., Zountsas, B., Garganis, K., Rosenow, F., Hermsen, A., von Oertzen, T.J., Diepgen, T. L., Avanzini, G., Consortium, E., 2017. Histopathological Findings in Brain Tissue Obtained during Epilepsy Surgery. N. Engl. J. Med. 377, 1648-1656.

Boedhoe, P.S., Schmaal, L., Abe, Y., Ameis, S.H., Arnold, P.D., Batistuzzo, M.C., Benedetti, F., Beucke, J.C., Bollettini, I., Bose, A., Brem, S., Calvo, A., Cheng, Y., Cho, K.I., Dallaspezia, S., Denys, D., Fitzgerald, K.D., Fouche, J.P., Gimenez, M., Gruner, P., Hanna, G.L., Hibar, D.P., Hoexter, M.Q., Hu, H., Huyser, C., Ikari, K., Jahanshad, N., Kathmann, N., Kaufmann, C., Koch, K., Kwon, J.S., Lazaro, L., Liu, Y., Lochner, C., Marsh, R., Martinez-Zalacain, I., Mataix-Cols, D., Menchon, J.M., Minuzzi, L.,

Nakamae, T., Nakao, T., Narayanaswamy, J.C., Piras, F., Piras, F., Pittenger, C., Reddy, Y.C., Sato, J.R., Simpson, H.B., Soreni, N., Soriano-Mas, C., Spalletta, G., Stevens, M.C., Szeszko, P.R., Tolin, D.F., Venkatasubramanian, G., Walitza, S., Wang, Z., van Wingen, G.A., Xu, J., Xu, X., Yun, J.Y., Zhao, Q., Group, E.O.W., Thompson, P.M., Stein, D.J., van den Heuvel, O.A., 2017. Distinct Subcortical Volume Alterations in Pediatric and Adult OCD: A Worldwide Meta- and Mega-Analysis. Am. J. Psychiatry 174, 60-69.

Bonilha, L., Edwards, J.C., Kinsman, S.L., Morgan, P.S., Fridriksson, J., Rorden, C., Rumboldt, Z., Roberts, D.R., Eckert, M.A., Halford, J.J., 2010a. Extrahippocampal gray matter loss and hippocampal deafferentation in patients with temporal lobe epilepsy. Epilepsia 51, 519–528.

Bonilha, L., Elm, J.J., Edwards, J.C., Morgan, P.S., Hicks, C., Lozar, C., Rumboldt, Z., Roberts, D.R., Rorden, C., Eckert, M.A., 2010b. How common is brain atrophy in patients with medial temporal lobe epilepsy? Epilepsia 51, 1774–1779.

Bonilha, L., Jensen, J.H., Baker, N., Breedlove, J., Nesland, T., Lin, J.J., Drane, D.L., Saindane, A.M., Binder, J.R., Kuzniecky, R.I., 2015. The brain connectome as a personalized biomarker of seizure outcomes after temporal lobectomy. Neurology.

Bonilha, L., Keller, S.S., 2015. Quantitative MRI in refractory temporal lobe epilepsy: relationship with surgical outcomes. Quant Imaging Med. Surg. 5, 204–224.

Bonilha, L., Kobayashi, E., Rorden, C., Cendes, F., Li, L.M., 2003. Medial temporal lobe atrophy in patients with refractory temporal lobe epilepsy. J. Neurol. Neurosurg. Psychiatry 74, 1627–1630.

Caligiuri, M.E., Labate, A., Cherubini, A., Mumoli, L., Ferlazzo, E., Aguglia, U., Quattrone, A., Gambardella, A., 2016. Integrity of the corpus callosum in patients with benign temporal lobe epilepsy. Epilepsia 57, 590–596.

Concha, L., Kim, H., Bernasconi, A., Bernhardt, B.C., Bernasconi, N., 2012. Spatial patterns of water diffusion along white matter tracts in temporal lobe epilepsy. Neurology 79, 455–462.

Coudray, N., Ocampo, P.S., Sakellaropoulos, T., Narula, N., Snuderl, M., Fenyo, D., Moreira, A.L., Razavian, N., Tsirigos, A., 2018. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. Nat. Med. 24, 1559–1567.

Davoodi-Bojd, E., Elisevich, K.V., Schwalb, J., Air, E., Soltanian-Zadeh, H., 2016. TLE lateralization using whole brain structural connectivity. Conf Proc IEEE Eng Med Biol Soc 2016, 1103–1106.

Del Gaizo, J., Mofrad, N., Jensen, J.H., Clark, D., Glenn, R., Helpern, J., Bonilha, L., 2017. Using machine learning to classify temporal lobe epilepsy based on diffusion MRI. Brain Behav 7, e00801.

Engel Jr., J., Thompson, P.M., Stern, J.M., Staba, R.J., Bragin, A., Mody, I., 2013. Connectomics and epilepsy. Curr. Opin. Neurol. 26, 186–194.

Faria, A.V., Zhang, J., Oishi, K., Li, X., Jiang, H., Akhter, K., Hermoye, L., Lee, S.K., Hoon, A., Stashinko, E., Miller, M.I., van Zijl, P.C., Mori, S., 2010. Atlas-based analysis of neurodevelopment from infancy to adulthood using diffusion tensor imaging and applications for automated abnormality detection. Neuroimage 52, 415–428.

Fischl, B., 2012. FreeSurfer. Neuroimage 62, 774–781.

Focke, N.K., Yogarajah, M., Bonelli, S.B., Bartlett, P.A., Symms, M.R., Duncan, J.S., 2008. Voxel-based diffusion tensor imaging in patients with mesial temporal lobe epilepsy and hippocampal sclerosis. Neuroimage 40, 728–737.

Focke, N.K., Yogarajah, M., Symms, M.R., Gruber, O., Paulus, W., Duncan, J.S., 2012. Automated MR image classification in temporal lobe epilepsy. Neuroimage 59, 356–362.

Fortin, J.P., Parker, D., Tunc, B., Watanabe, T., Elliott, M.A., Ruparel, K., Roalf, D.R., Satterthwaite, T.D., Gur, R.C., Gur, R.E., Schultz, R.T., Verma, R., Shinohara, R.T., 2017. Harmonization of multi-site diffusion tensor imaging data. Neuroimage 161, 149–170.

Gleichgerrcht, E., Bonilha, L., 2017. Structural brain network architecture and personalized medicine in epilepsy. Exp. Rev. Precision Med. Drug Dev. 2, 229–237.

Gleichgerrcht, E., Keller, S.S., Drane, D.L., Munsell, B.C., Davis, K.A., Kaestner, E., Weber, B., Krantz, S., Vandergrift, W.A., Edwards, J.C., McDonald, C.R., Kuzniecky, R., Bonilha, L., 2020. Temporal Lobe Epilepsy Surgical Outcomes Can Be Inferred Based on Structural Connectome Hubs: A Machine Learning Study. Ann. Neurol. 88, 970–983.

Gleichgerrcht, E., Kocher, M., Bonilha, L., 2015. Connectomics and graph theory analyses: Novel insights into network abnormalities in epilepsy. Epilepsia 56, 1660–1668.

Gleichgerrcht, E., Munsell, B., Bhatia, S., Vandergrift 3rd, W.A., Rorden, C., McDonald, C., Edwards, J., Kuzniecky, R., Bonilha, L., 2018. Deep learning applied to whole-brain connectome to determine seizure control after epilepsy surgery. Epilepsia 59, 1643–1654.

Gulshan, V., Peng, L., Coram, M., Stumpe, M.C., Wu, D., Narayanaswamy, A., Venugopalan, S., Widner, K., Madams, T., Cuadros, J., Kim, R., Raman, R., Nelson, P. C., Mega, J.L., Webster, D.R., 2016. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. JAMA 316, 2402–2410.

Haenssle, H.A., Fink, C., Schneiderbauer, R., Toberer, F., Buhl, T., Blum, A., Kalloo, A., Hassen, A.B.H., Thomas, L., Enk, A., Uhlmann, L., level, I., level, I.I.G., Alt, C., Arenbergerova, M., Bakos, R., Baltzer, A., Bertlich, I., Blum, A., Bokor-Billmann, T., Bowling, J., Braghiroli, N., Braun, R., Buder-Bakhaya, K., Buhl, T., Cabo, H., Cabrijan, L., Cevic, N., Classen, A., Deltgen, D., Fink, C., Georgieva, I., Hakim-Meibodi, L.E., Hanner, S., Hartmann, F., Hartmann, J., Haus, G., Hoxha, E., Karls, R., Koga, H., Kreusch, J., Lallas, A., Majenka, P., Marghoob, A., Massone, C., Mekokishvili, L., Mestel, D., Meyer, V., Neuberger, A., Nielsen, K., Oliviero, M., Pampena, R., Paoli, J., Pawlik, E., Rao, B., Rendon, A., Russo, T., Sadek, A., Samhaber, K., Schneiderbauer, R., Schweizer, A., Toberer, F., Trennheuser, L., Vlahova, L., Wald, A., Winkler, J., Wolbing, P., Zalaudek, I., 2018. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. Ann. Oncol. 29, 1836-1842.

Hatton, S.N., Huynh, K.H., Bonilha, L., Abela, E., Alhusaini, S., Altmann, A., Alvim, M.K. M., Balachandra, A.R., Bartolini, E., Bender, B., Bernasconi, N., Bernasconi, A., Bernhardt, B., Bargallo, N., Caldairou, B., Caligiuri, M.E., Carr, S.J.A., Cavalleri, G.L., Cendes, F., Concha, L., Davoodi-Bojd, E., Desmond, P.M., Devinsky, O., Doherty, C. P., Domin, M., Duncan, J.S., Focke, N.K., Foley, S.F., Gambardella, A., Gleichgerrcht, E., Guerrini, R., Hamandi, K., Ishikawa, A., Keller, S.S., Kochunov, P. V., Kotikalapudi, R., Kreilkamp, B.A.K., Kwan, P., Labate, A., Langner, S., Lenge, M., Liu, M., Lui, E., Martin, P., Mascalchi, M., Moreira, J.C.V., Morita-Sherman, M.E., O'Brien, T.J., Pardoe, H.R., Pariente, J.C., Ribeiro, L.F., Richardson, M.P., Rocha, C. S., Rodriguez-Cruces, R., Rosenow, F., Severino, M., Sinclair, B., Soltanian-Zadeh, H., Striano, P., Taylor, P.N., Thomas, R.H., Tortora, D., Velakoulis, D., Vezzani, A., Vivash, L., von Podewils, F., Vos, S.B., Weber, B., Winston, G.P., Yasuda, C.L., Zhu, A.H., Thompson, P.M., Whelan, C.D., Jahanshad, N., Sisodiya, S.M., McDonald, C.R., 2020. White matter abnormalities across different epilepsy syndromes in adults: an ENIGMA-Epilepsy study. Brain 143, 2454–2473.

Hazlett, H.C., Gu, H., Munsell, B.C., Kim, S.H., Styner, M., Wolff, J.J., Elison, J.T., Swanson, M.R., Zhu, H., Botteron, K.N., Collins, D.L., Constantino, J.N., Dager, S.R., Estes, A.M., Evans, A.C., Fonov, V.S., Gerig, G., Kostopoulos, P., McKinstry, R.C., Pandey, J., Paterson, S., Pruett, J.R., Schultz, R.T., Shaw, D.W., Zwaigenbaum, L., Piven, J., IBIS Network, Clinical Sites, Data Coordinating Center, Image Processing Core, Statistical Analysis, 2017. Early brain development in infants at high risk for autism spectrum disorder. Nature 542, 348–351. https://doi.org/10.1038/nature21369. PMID: 28202961.

Hibar, D.P., Adams, H.H.H., Jahanshad, N., Chauhan, G., Stein, J.L., Hofer, E., Renteria, M.E., Bis, J.C., Arias-Vasquez, A., Ikram, M.K., Desrivieres, S., Vernooij, M.W., Abramovic, L., Alhusaini, S., Amin, N., Andersson, M., Arfanakis, K., Aribisala, B.S., Armstrong, N.J., Athanasiu, L., Axelsson, T., Beecham, A.H., Beiser, A., Bernard, M., Blanton, S.H., Bohlken, M.M., Boks, M.P., Bralten, J., Brickman, A.M., Carmichael, O., Chakravarty, M.M., Chen, Q., Ching, C.R.K., Chouraki, V., Cuellar-Partida, G., Crivello, F., Den Braber, A., Doan, N.T., Ehrlich, S., Giddaluru, S., Goldman, A.L., Gottesman, R.F., Grimm, O., Griswold, M.E., Guadalupe, T., Gutman, B.A., Hass, J., Haukvik, U.K., Hoehn, D., Holmes, A.J., Hoogman, M., Janowitz, D., Jia, T., Jorgensen, K.N., Karbalai, N., Kasperaviciute, D., Kim, S., Klein, M., Kraemer, B., Lee, P.H., Liewald, D.C.M., Lopez, L.M., Luciano, M., Macare, C., Marquand, A.F., Matarin, M., Mather, K.A., Mattheisen, M., McKay, D.R., Milaneschi, Y., Munoz Maniega, S., Nho, K., Nugent, A.C., Nyquist, P., Loohuis, L.M.O., Oosterlaan, J., Papmeyer, M., Pirpamer, L., Putz, B., Ramasamy, A., Richards, J.S., Risacher, S.L., Roiz-Santianez, R., Rommelse, N., Ropele, S., Rose, E.J., Royle, N.A., Rundek, T., Samann, P.G., Saremi, A., Satizabal, C.L., Schmaal, L., Schork, A.J., Shen, L., Shin, J., Shumskaya, E., Smith, A.V., Sprooten, E., Strike, L.T., Teumer, A., Tordesillas-Gutierrez, D., Toro, R., Trabzuni, D., Trompet, S., Vaidya, D., Van der Grond, J., Van der Lee, S.J., Van der Meer, D., Van Donkelaar, M.M.J., Van Eijk, K.R., Van Erp, T.G. M., Van Rooij, D., Walton, E., Westlye, L.T., Whelan, C.D., Windham, B.G., Winkler, A.M., Wittfeld, K., Woldehawariat, G., Wolf, C., Wolfers, T., Yanek, L.R., Yang, J., Zijdenbos, A., Zwiers, M.P., Agartz, I., Almasy, L., Ames, D., Amouyel, P., Andreassen, O.A., Arepalli, S., Assareh, A.A., Barral, S., Bastin, M.E., Becker, D.M., Becker, J.T., Bennett, D.A., Blangero, J., van Bokhoven, H., Boomsma, D.I., Brodaty, H., Brouwer, R.M., Brunner, H.G., Buckner, R.L., Buitelaar, J.K., Bulayeva, K.B., Cahn, W., Calhoun, V.D., Cannon, D.M., Cavalleri, G.L., Cheng, C.Y., Cichon, S., Cookson, M.R., Corvin, A., Crespo-Facorro, B., Curran, J.E., Czisch, M., Dale, A.M., Davies, G.E., De Craen, A.J.M., De Geus, E.J.C., De Jager, P.L., De Zubicaray, G.I., Deary, I.J., Debette, S., DeCarli, C., Delanty, N., Depondt, C., DeStefano, A., Dillman, A., Djurovic, S., Donohoe, G., Drevets, W.C., Duggirala, R., Dyer, T.D., Enzinger, C., Erk, S., Espeseth, T., Fedko, I.O., Fernandez, G., Ferrucci, L., Fisher, S.E., Fleischman, D.A., Ford, I., Fornage, M., Foroud, T.M., Fox, P.T., Francks, C., Fukunaga, M., Gibbs, J.R., Glahn, D.C., Gollub, R.L., Goring, H.H.H., Green, R.C., Gruber, O., Gudnason, V., Guelfi, S., Haberg, A.K., Hansell, N.K., Hardy, J., Hartman, C.A., Hashimoto, R., Hegenscheid, K., Heinz, A., Le Hellard, S., Hernandez, D.G., Heslenfeld, D.J., Ho, B. C., Hoekstra, P.J., Hoffmann, W., Hofman, A., Holsboer, F., Homuth, G., Hosten, N., Hottenga, J.J., Huentelman, M., Hulshoff Pol, H.E., Ikeda, M., Jack, C.R., Jr., Jenkinson, M., Johnson, R., Jonsson, E.G., Jukema, J.W., Kahn, R.S., Kanai, R., Kloszewska, I., Knopman, D.S., Kochunov, P., Kwok, J.B., Lawrie, S.M., Lemaitre, H., Liu, X., Longo, D.L., Lopez, O.L., Lovestone, S., Martinez, O., Martinot, J.L., Mattay, V.S., McDonald, C., McIntosh, A.M., McMahon, F.J., McMahon, K.L., Mecocci, P., Melle, I., Meyer-Lindenberg, A., Mohnke, S., Montgomery, G.W., Morris, D.W., Mosley, T.H., Muhleisen, T.W., Muller-Myhsok, B., Nalls, M.A., Nauck, M., Nichols, T.E., Niessen, W.J., Nothen, M.M., Nyberg, L., Ohi, K., Olvera, R.L., Ophoff, R.A., Pandolfo, M., Paus, T., Pausova, Z., Penninx, B., Pike, G.B., Potkin, S.G., Psaty, B.M., Reppermund, S., Rietschel, M., Roffman, J.L., Romanczuk-Seiferth, N., Rotter, J.I., Ryten, M., Sacco, R.L., Sachdev, P.S., Saykin, A.J., Schmidt, R., Schmidt, H., Schofield, P.R., Sigursson, S., Simmons, A., Singleton, A., Sisodiya, S.M., Smith, C., Smoller, J.W., Soininen, H., Steen, V.M., Stott, D.J., Sussmann, J.E., Thalamuthu, A., Toga, A.W., Traynor, B.J., Troncoso, J., Tsolaki, M., Tzourio, C., Uitterlinden, A.G., Hernandez, M.C.V., Van der Brug, M., van der Lugt, A., van der Wee, N.J.A., Van Haren, N.E.M., van 't Ent, D., Van Tol, M.J., Vardarajan, B.N., Vellas, B., Veltman, D. J., Volzke, H., Walter, H., Wardlaw, J.M., Wassink, T.H., Weale, M.E., Weinberger, D.R., Weiner, M.W., Wen, W., Westman, E., White, T., Wong, T.Y., Wright, C.B., Zielke, R.H., Zonderman, A.B., Martin, N.G., Van Duijn, C.M., Wright, M.J., Longstreth, W.T., Schumann, G., Grabe, H.J., Franke, B., Launer, L.J., Medland, S.E., Seshadri, S., Thompson, P.M., Ikram, M.A., 2017. Novel genetic loci associated with hippocampal volume. Nat. Commun. 8, 13624.

Hibar, D.P., Stein, J.L., Renteria, M.E., Arias-Vasquez, A., Desrivieres, S., Jahanshad, N., Toro, R., Wittfeld, K., Abramovic, L., Andersson, M., Aribisala, B.S., Armstrong, N.J.,

Bernard, M., Bohlken, M.M., Boks, M.P., Bralten, J., Brown, A.A., Chakravarty, M.M., Chen, Q., Ching, C.R., Cuellar-Partida, G., den Braber, A., Giddaluru, S., Goldman, A. L., Grimm, O., Guadalupe, T., Hass, J., Woldehawariat, G., Holmes, A.J., Hoogman, M., Janowitz, D., Jia, T., Kim, S., Klein, M., Kraemer, B., Lee, P.H., Olde Loohuis, L. M., Luciano, M., Macare, C., Mather, K.A., Mattheisen, M., Milaneschi, Y., Nho, K., Papmeyer, M., Ramasamy, A., Risacher, S.L., Roiz-Santianez, R., Rose, E.J., Salami, A., Samann, P.G., Schmaal, L., Schork, A.J., Shin, J., Strike, L.T., Teumer, A., van Donkelaar, M.M., van Eijk, K.R., Walters, R.K., Westlye, L.T., Whelan, C.D., Winkler, A.M., Zwiers, M.P., Alhusaini, S., Athanasiu, L., Ehrlich, S., Hakobjan, M.M., Hartberg, C.B., Haukvik, U.K., Heister, A.J., Hoehn, D., Kasperaviciute, D., Liewald, D.C., Lopez, L.M., Makkinje, R.R., Matarin, M., Naber, M.A., McKay, D.R., Needham, M., Nugent, A.C., Putz, B., Royle, N.A., Shen, L., Sprooten, E., Trabzuni, D., van der Marel, S.S., van Hulzen, K.J., Walton, E., Wolf, C., Almasy, L., Ames, D., Arepalli, S., Assaraf, A.A., Bastin, M.E., Brodaty, H., Bulayeva, K.B., Carless, M.A., Cichon, S., Corvin, A., Curran, J.E., Czisch, M., de Zubicaray, G.I., Dillman, A., Duggirala, R., Dyer, T.D., Erk, S., Fedko, I.O., Ferrucci, L., Foroud, T.M., Fox, P.T., Fukunaga, M., Gibbs, J.R., Goring, H.H., Green, R.C., Guelfi, S., Hansell, N.K., Hartman, C.A., Hegenscheid, K., Heinz, A., Hernandez, D.G., Heslenfeld, D.J., Hoekstra, P.J., Holsboer, F., Homuth, G., Hottenga, J.J., Ikeda, M., Jack, C.R., Jr., Jenkinson, M., Johnson, R., Kanai, R., Keil, M., Kent, J.W., Jr., Kochunov, P., Kwok, J.B., Lawrie, S. M., Liu, X., Longo, D.L., McMahon, K.L., Meisenzahl, E., Melle, I., Mohnke, S., Montgomery, G.W., Mostert, J.C., Muhleisen, T.W., Nalls, M.A., Nichols, T.E., Nilsson, L.G., Nothen, M.M., Ohi, K., Olvera, R.L., Perez-Iglesias, R., Pike, G.B., Potkin, S.G., Reinvang, I., Reppermund, S., Rietschel, M., Romanczuk-Seiferth, N., Rosen, G.D., Rujescu, D., Schnell, K., Schofield, P.R., Smith, C., Steen, V.M., Sussmann, J.E., Thalamuthu, A., Toga, A.W., Traynor, B.J., Troncoso, J., Turner, J. A., Valdes Hernandez, M.C., van 't Ent, D., van der Brug, M., van der Wee, N.J., van Tol, M.J., Veltman, D.J., Wassink, T.H., Westman, E., Zielke, R.H., Zonderman, A.B., Ashbrook, D.G., Hager, R., Lu, L., McMahon, F.J., Morris, D.W., Williams, R.W., Brunner, H.G., Buckner, R.L., Buitelaar, J.K., Cahn, W., Calhoun, V.D., Cavalleri, G. L., Crespo-Facorro, B., Dale, A.M., Davies, G.E., Delanty, N., Depondt, C., Djurovic, S., Drevets, W.C., Espeseth, T., Gollub, R.L., Ho, B.C., Hoffmann, W., Hosten, N., Kahn, R.S., Le Hellard, S., Meyer-Lindenberg, A., Muller-Myhsok, B., Nauck, M., Nyberg, L., Pandolfo, M., Penninx, B.W., Roffman, J.L., Sisodiya, S.M., Smoller, J.W., van Bokhoven, H., van Haren, N.E., Volzke, H., Walter, H., Weiner, M.W., Wen, W., White, T., Agartz, I., Andreassen, O.A., Blangero, J., Boomsma, D.I., Brouwer, R.M., Cannon, D.M., Cookson, M.R., de Geus, E.J., Deary, I.J., Donohoe, G., Fernandez, G., Fisher, S.E., Francks, C., Glahn, D.C., Grabe, H.J., Gruber, O., Hardy, J., Hashimoto, R., Hulshoff Pol, H.E., Jonsson, E.G., Kloszewska, I., Lovestone, S., Mattay, V.S., Mecocci, P., McDonald, C., McIntosh, A.M., Ophoff, R.A., Paus, T., Pausova, Z., Ryten, M., Sachdev, P.S., Saykin, A.J., Simmons, A., Singleton, A., Soininen, H., Wardlaw, J.M., Weale, M.E., Weinberger, D.R., Adams, H.H., Launer, L.J., Seiler, S., Schmidt, R., Chauhan, G., Satizabal, C.L., Becker, J.T., Yanek, L., van der Lee, S.J., Ebling, M., Fischl, B., Longstreth, W.T., Jr., Greve, D., Schmidt, H., Nyquist, P., Vinke, L.N., van Duijn, C.M., Xue, L., Mazoyer, B., Bis, J.C., Gudnason, V., Seshadri, S., Ikram, M.A., Alzheimer's Disease Neuroimaging, I., Consortium, C., Epigen, Imagen, Sys, Martin, N.G., Wright, M.J., Schumann, G., Franke, B., Thompson, P.M., Medland, S.E., 2015. Common genetic variants influence human subcortical brain structures. Nature 520, 224-229.

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

Hoogman, M., Bralten, J., Hibar, D.P., Mennes, M., Zwiers, M.P., Schweren, L.S.J., van Hulzen, K.J.E., Medland, S.E., Shumskaya, E., Jahanshad, N., Zeeuw, P., Szekely, E., Sudre, G., Wolfers, T., Onnink, A.M.H., Dammers, J.T., Mostert, J.C., Vives-Gilabert, Y., Kohls, G., Oberwelland, E., Seitz, J., Schulte-Ruther, M., Ambrosino, S., Doyle, A.E., Hovik, M.F., Dramsdahl, M., Tamm, L., van Erp, T.G.M., Dale, A., Schork, A., Conzelmann, A., Zierhut, K., Baur, R., McCarthy, H., Yoncheva, Y.N., Cubillo, A., Chantiluke, K., Mehta, M.A., Paloyelis, Y., Hohmann, S., Baumeister, S., Bramati, I., Mattos, P., Tovar-Moll, F., Douglas, P., Banaschewski, T., Brandeis, D., Kuntsi, J., Asherson, P., Rubia, K., Kelly, C., Martino, A.D., Milham, M.P., Castellanos, F.X., Frodl, T., Zentis, M., Lesch, K.P., Reif, A., Pauli, P., Jernigan, T.L., Haavik, J., Plessen, K.J., Lundervold, A.J., Hugdahl, K., Seidman, L.J., Biederman, J., Rommelse, N., Heslenfeld, D.J., Hartman, C.A., Hoekstra, P.J., Oosterlaan, J., Polier, G.V., Konrad, K., Vilarroya, O., Ramos-Quiroga, J.A., Soliva, J.C., Durston, S., Buitelaar, J.K., Faraone, S.V., Shaw, P., Thompson, P.M., Franke, B., 2017. Subcortical brain volume differences in participants with attention deficit hyperactivity disorder in children and adults: a cross-sectional mega-analysis. Lancet Psychiatry 4, 310–319.

Johnson, W.E., Li, C., Rabinovic, A., 2007. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 8, 118–127.

Kamiya, K., Amemiya, S., Suzuki, Y., Kunii, N., Kawai, K., Mori, H., Kunimatsu, A., Saito, N., Aoki, S., Ohtomo, K., 2016. Machine Learning of DTI Structural Brain Connectomes for Lateralization of Temporal Lobe Epilepsy. Magn Reson Med Sci 15, 121–129.

Lariviere, S., Rodriguez-Cruces, R., Royer, J., Caligiuri, M.E., Gambardella, A., Concha, L., Keller, S.S., Cendes, F., Yasuda, C., Bonilha, L., Gleichgerrcht, E., Focke, N.K., Domin, M., von Podewills, F., Langner, S., Rummel, C., Wiest, R., Martin, P., Kotikalapudi, R., O'Brien, T.J., Sinclair, B., Vivash, L., Desmond, P.M., Alhusaini, S., Doherty, C.P., Cavalleri, G.L., Delanty, N., Kalviainen, R., Jackson, G.D., Kowalczyk, M., Mascalchi, M., Semmelroch, M., Thomas, R.H., Soltanian-Zadeh, H., Davoodi-Bojd, E., Zhang, J., Lenge, M., Guerrini, R., Bartolini, E., Hamandi, K., Foley, S., Weber, B., Depondt, C., Absil, J., Carr, S.J.A., Abela, E., Richardson, M.P., Devinsky, O., Severino, M., Striano, P., Tortora, D., Hatton, S.N., Vos, S.B., Duncan, J.S., Whelan, C.D., Thompson, P.M., Sisodiya, S.M., Bernasconi, A., Labate, A., McDonald,

C.R., Bernasconi, N., Bernhardt, B.C., 2020. Network-based atrophy modeling in the common epilepsies: A worldwide ENIGMA study. Sci Adv 6.

Lee, J.S., Lee, D.S., Kim, S.K., Lee, S.K., Chung, J.K., Lee, M.C., Park, K.S., 2000. Localization of epileptogenic zones in F-18 FDG brain PET of patients with temporal lobe epilepsy using artificial neural network. IEEE Trans. Med. Imaging 19, 347–355.

Lin, J.J., Salamon, N., Lee, A.D., Dutton, R.A., Geaga, J.A., Hayashi, K.M., Luders, E., Toga, A.W., Engel Jr., J., Thompson, P.M., 2007. Reduced neocortical thickness and complexity mapped in mesial temporal lobe epilepsy with hippocampal sclerosis. Cereb. Cortex 17, 2007–2018.

Liu, M., Bernhardt, B.C., Bernasconi, A., Bernasconi, N., 2016. Gray matter structural compromise is equally distributed in left and right temporal lobe epilepsy. Hum. Brain Mapp. 37, 515–524.

Lopes, R., Steinling, M., Szurhaj, W., Maouche, S., Dubois, P., Betrouni, N., 2010. Fractal features for localization of temporal lobe epileptic foci using SPECT imaging. Comput. Biol. Med. 40, 469–477.

Margerison, J.H., Corsellis, J.A., 1966. Epilepsy and the temporal lobes. A clinical, electroencephalographic and neuropathological study of the brain in epilepsy, with particular reference to the temporal lobes. Brain 89, 499–530.

McDonald, C.R., Ahmadi, M.E., Hagler, D.J., Tecoma, E.S., Iragui, V.J., Gharapetian, L., Dale, A.M., Halgren, E., 2008a. Diffusion tensor imaging correlates of memory and language impairments in temporal lobe epilepsy. Neurology 71, 1869–1876.

McDonald, C.R., Hagler Jr., D.J., Ahmadi, M.E., Tecoma, E., Iragui, V., Gharapetian, L., Dale, A.M., Halgren, E., 2008b. Regional neocortical thinning in mesial temporal lobe epilepsy. Epilepsia 49, 794–803.

Munsell, B.C., Wee, C.Y., Keller, S.S., Weber, B., Elger, C., da Silva, L.A., 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.

Munsell, B.C., Gleichgerrcht, E., Hofesmann, E., Delgaizo, J., McDonald, C.R., Marebwa, B., Styner, M.A., Fridriksson, J., Rorden, C., Focke, N.K., Gilmore, J.H., Bonilha, L., 2020 Nov 1. Personalized connectome fingerprints: Their importance in cognition from childhood to adult years. Neuroimage 221, 117122. https://doi.org/10.1016/j.neuroimage.2020.117122. Epub 2020 Jul 4. PMID: 32634596.

Poldrack, R.A., Huckins, G., Varoquaux, G., 2019. Establishment of Best Practices for Evidence for Prediction: A Review. JAMA. Psychiatry.

Richardson, M.P., 2012. Large scale brain models of epilepsy: dynamics meets connectomics. J. Neurol. Neurosurg. Psychiatry 83, 1238–1248.

Rodriguez-Cruces, R., Bernhardt, B.C., Concha, L., 2020. Multidimensional associations between cognition and connectome organization in temporal lobe epilepsy. Neuroimage 213, 116706.

Rudie, J.D., Colby, J.B., Salamon, N., 2015. Machine learning classification of mesial temporal sclerosis in epilepsy patients. Epilepsy Res. 117, 63–69.

Schmaal, L., Hibar, D.P., Samann, P.G., Hall, G.B., Baune, B.T., Jahanshad, N., Cheung, J. W., van Erp, T.G.M., Bos, D., Ikram, M.A., Vernooij, M.W., Niessen, W.J., Tiemeier, H., Hofman, A., Wittfeld, K., Grabe, H.J., Janowitz, D., Bulow, R., Selonke, M., Volzke, H., Grotegerd, D., Dannlowski, U., Arolt, V., Opel, N., Heindel, W., Kugel, H., Hoehn, D., Czisch, M., Couvy-Duchesne, B., Renteria, M.E., Strike, L.T., Wright, M.J., Mills, N.T., de Zubicaray, G.I., McMahon, K.L., Medland, S. E., Martin, N.G., Gillespie, N.A., Goya-Maldonado, R., Gruber, O., Kramer, B., Hatton, S.N., Lagopoulos, J., Hickie, I.B., Frodl, T., Carballedo, A., Frey, E.M., van Velzen, L.S., Penninx, B., van Tol, M.J., van der Wee, N.J., Davey, C.G., Harrison, B. J., Mwangi, B., Cao, B., Soares, J.C., Veer, I.M., Walter, H., Schoepf, D., Zurowski, B., Konrad, C., Schramm, E., Normann, C., Schnell, K., Sacchet, M.D., Gotlib, I.H., MacQueen, G.M., Godlewska, B.R., Nickson, T., McIntosh, A.M., Papmeyer, M., Whalley, H.C., Hall, J., Sussmann, J.E., Li, M., Walter, M., Aftanas, L., Brack, I., Bokhan, N.A., Thompson, P.M., Veltman, D.J., 2017. Cortical abnormalities in adults and adolescents with major depression based on brain scans from 20 cohorts worldwide in the ENIGMA Major Depressive Disorder Working Group. Mol. Psychiatry 22, 900–909.

Stein, J.L., Medland, S.E., Vasquez, A.A., Hibar, D.P., Senstad, R.E., Winkler, A.M., Toro, R., Appel, K., Bartecek, R., Bergmann, O., Bernard, M., Brown, A.A., Cannon, D.M., Chakravarty, M.M., Christoforou, A., Domin, M., Grimm, O., Hollinshead, M., Holmes, A.J., Homuth, G., Hottenga, J.J., Langan, C., Lopez, L.M., Hansell, N.K., Hwang, K.S., Kim, S., Laje, G., Lee, P.H., Liu, X., Loth, E., Lourdusamy, A., Mattingsdal, M., Mohnke, S., Maniega, S.M., Nho, K., Nugent, A.C., O'Brien, C., Papmeyer, M., Putz, B., Ramasamy, A., Rasmussen, J., Rijpkema, M., Risacher, S.L., Roddey, J.C., Rose, E.J., Ryten, M., Shen, L., Sprooten, E., Strengman, E., Teumer, A., Trabzuni, D., Turner, J., van Eijk, K., van Erp, T.G., van Tol, M.J., Wittfeld, K., Wolf, C., Woudstra, S., Aleman, A., Alhusaini, S., Almasy, L., Binder, E.B., Brohawn, D.G., Cantor, R.M., Carless, M.A., Corvin, A., Czisch, M., Curran, J.E., Davies, G., de Almeida, M.A., Delanty, N., Depondt, C., Duggirala, R., Dyer, T.D., Erk, S., Fagerness, J., Fox, P.T., Freimer, N.B., Gill, M., Goring, H.H., Hagler, D.J., Hoehn, D., Holsboer, F., Hoogman, M., Hosten, N., Jahanshad, N., Johnson, M.P., Kasperaviciute, D., Kent, J.W., Jr., Kochunov, P., Lancaster, J.L., Lawrie, S.M., Liewald, D.C., Mandl, R., Matarin, M., Mattheisen, M., Meisenzahl, E., Melle, I., Moses, E.K., Muhleisen, T.W., Nauck, M., Nothen, M.M., Olvera, R.L., Pandolfo, M., Pike, G.B., Puls, R., Reinvang, I., Renteria, M.E., Rietschel, M., Roffman, J.L., Royle, N.A., Rujescu, D., Savitz, J., Schnack, H.G., Schnell, K., Seiferth, N., Smith, C., Steen, V.M., Valdes Hernandez, M. C., Van den Heuvel, M., van der Wee, N.J., Van Haren, N.E., Veltman, J.A., Volzke, H., Walker, R., Westlye, L.T., Whelan, C.D., Agartz, I., Boomsma, D.I., Cavalleri, G.L., Dale, A.M., Djurovic, S., Drevets, W.C., Hagoort, P., Hall, J., Heinz, A., Jack, C.R., Jr., Foroud, T.M., Le Hellard, S., Macciardi, F., Montgomery, G.W., Poline, J.B., Porteous, D.J., Sisodiya, S.M., Starr, J.M., Sussmann, J., Toga, A.W., Veltman, D.J., Walter, H., Weiner, M.W., Alzheimer's Disease Neuroimaging, I., Consortium, E., Consortium, I., Saguenay Youth Study, G., Bis, J.C., Ikram, M.A., Smith, A.V., Gudnason, V., Tzourio, C., Vernooij, M.W., Launer, L.J., DeCarli, C., Seshadri, S.,

Cohorts for, H., Aging Research in Genomic Epidemiology, C., Andreassen, O.A., Apostolova, L.G., Bastin, M.E., Blangero, J., Brunner, H.G., Buckner, R.L., Cichon, S., Coppola, G., de Zubicaray, G.I., Deary, I.J., Donohoe, G., de Geus, E.J., Espeseth, T., Fernandez, G., Glahn, D.C., Grabe, H.J., Hardy, J., Hulshoff Pol, H.E., Jenkinson, M., Kahn, R.S., McDonald, C., McIntosh, A.M., McMahon, F.J., McMahon, K.L., Meyer-Lindenberg, A., Morris, D.W., Muller-Myhsok, B., Nichols, T.E., Ophoff, R.A., Paus, T., Pausova, Z., Penninx, B.W., Potkin, S.G., Samann, P.G., Saykin, A.J., Schumann, G., Smoller, J.W., Wardlaw, J.M., Weale, M.E., Martin, N.G., Franke, B., Wright, M. J., Thompson, P.M., Enhancing Neuro Imaging Genetics through Meta-Analysis, C., 2012. Identification of common variants associated with human hippocampal and intracranial volumes. Nat. Genet. 44, 552-561.

Tavakol, S., Royer, J., Lowe, A.J., Bonilha, L., Tracy, J.I., Jackson, G.D., Duncan, J.S., Bernasconi, A., Bernasconi, N., Bernhardt, B.C., 2019. Neuroimaging and connectomics of drug-resistant epilepsy at multiple scales: From focal lesions to macroscale networks. Epilepsia 60, 593–604.

Tellez-Zenteno, J.F., Hernandez-Ronquillo, L., 2012. A review of the epidemiology of temporal lobe epilepsy. Epilepsy Res Treat 2012, 630853.

Thompson, P.M., Jahanshad, N., Ching, C.R.K., Salminen, L.E., Thomopoulos, S.I., Bright, J., Baune, B.T., Bertolin, S., Bralten, J., Bruin, W.B., Bulow, R., Chen, J., Chye, Y., Dannlowski, U., de Kovel, C.G.F., Donohoe, G., Eyler, L.T., Faraone, S.V., Favre, P., Filippi, C.A., Frodl, T., Garijo, D., Gil, Y., Grabe, H.J., Grasby, K.L., Hajek, T., Han, L. K.M., Hatton, S.N., Hilbert, K., Ho, T.C., Holleran, L., Homuth, G., Hosten, N., Houenou, J., Ivanov, I., Jia, T., Kelly, S., Klein, M., Kwon, J.S., Laansma, M.A., Leerssen, J., Lueken, U., Nunes, A., Neill, J.O., Opel, N., Piras, F., Piras, F., Postema, M.C., Pozzi, E., Shatokhina, N., Soriano-Mas, C., Spalletta, G., Sun, D., Teumer, A., Tilot, A.K., Tozzi, L., van der Merwe, C., Van Someren, E.J.W., van Wingen, G.A., Volzke, H., Walton, E., Wang, L., Winkler, A.M., Wittfeld, K., Wright, M.J., Yun, J.Y., Zhang, G., Zhang-James, Y., Adhikari, B.M., Agartz, I., Aghajani, M., Aleman, A., Althoff, R.R., Altmann, A., Andreassen, O.A., Baron, D.A., Bartnik-Olson, B.L., Marie Bas-Hoogendam, J., Baskin-Sommers, A.R., Bearden, C.E., Berner, L.A., Boedhoe, P. S.W., Brouwer, R.M., Buitelaar, J.K., Caeyenberghs, K., Cecil, C.A.M., Cohen, R.A., Cole, J.H., Conrod, P.J., De Brito, S.A., de Zwarte, S.M.C., Dennis, E.L., Desrivieres, S., Dima, D., Ehrlich, S., Esopenko, C., Fairchild, G., Fisher, S.E., Fouche, J.P., Francks, C., Frangou, S., Franke, B., Garavan, H.P., Glahn, D.C., Groenewold, N.A., Gurholt, T.P., Gutman, B.A., Hahn, T., Harding, I.H., Hernaus, D., Hibar, D.P., Hillary, F.G., Hoogman, M., Hulshoff Pol, H.E., Jalbrzikowski, M., Karkashadze, G. A., Klapwijk, E.T., Knickmeyer, R.C., Kochunov, P., Koerte, I.K., Kong, X.Z., Liew, S. L., Lin, A.P., Logue, M.W., Luders, E., Macciardi, F., Mackey, S., Mayer, A.R., McDonald, C.R., McMahon, A.B., Medland, S.E., Modinos, G., Morey, R.A., Mueller, S.C., Mukherjee, P., Namazova-Baranova, L., Nir, T.M., Olsen, A., Paschou, P., Pine, D.S., Pizzagalli, F., Renteria, M.E., Rohrer, J.D., Samann, P.G., Schmaal, L., Schumann, G., Shiroishi, M.S., Sisodiya, S.M., Smit, D.J.A., Sonderby, I.E., Stein, D. J., Stein, J.L., Tahmasian, M., Tate, D.F., Turner, J.A., van den Heuvel, O.A., van der Wee, N.J.A., van der Werf, Y.D., van Erp, T.G.M., van Haren, N.E.M., van Rooij, D.,

van Velzen, L.S., Veer, I.M., Veltman, D.J., Villalon-Reina, J.E., Walter, H., Whelan, C.D., Wilde, E.A., Zarei, M., Zelman, V., Consortium, E., 2020. ENIGMA and global neuroscience: A decade of large-scale studies of the brain in health and disease across more than 40 countries. Transl Psychiatry 10, 100.

Villalon-Reina, J.E., Martinez, K., Qu, X., Ching, C.R.K., Nir, T.M., Kothapalli, D., Corbin, C., Sun, D., Lin, A., Forsyth, J.K., Kushan, L., Vajdi, A., Jalbrzikowski, M., Hansen, L., Jonas, R.K., van Amelsvoort, T., Bakker, G., Kates, W.R., Antshel, K.M., Fremont, W., Campbell, L.E., McCabe, K.L., Daly, E., Gudbrandsen, M., Murphy, C.M., Murphy, D., Craig, M., Emanuel, B., McDonald-McGinn, D.M., Vorstman, J.A.S., Fiksinski, A.M., Koops, S., Ruparel, K., Roalf, D., Gur, R.E., Eric Schmitt, J., Simon, T.J., Goodrich-Hunsaker, N.J., Durdle, C.A., Doherty, J.L., Cunningham, A.C., van den Bree, M., Linden, D.E.J., Owen, M., Moss, H., Kelly, S., Donohoe, G., Murphy, K.C., Arango, C., Jahanshad, N., Thompson, P.M., Bearden, C.E., 2019. Altered white matter microstructure in 22q11.2 deletion syndrome: a multisite diffusion tensor imaging study. Mol Psychiatry.

Whelan, C.D., Altmann, A., Botia, J.A., Jahanshad, N., Hibar, D.P., Absil, J., Alhusaini, S., Alvim, M.K.M., Auvinen, P., Bartolini, E., Bergo, F.P.G., Bernardes, T., Blackmon, K., Braga, B., Caligiuri, M.E., Calvo, A., Carr, S.J., Chen, J., Chen, S., Cherubini, A., David, P., Domin, M., Foley, S., Franca, W., Haaker, G., Isaev, D., Keller, S.S., Kotikalapudi, R., Kowalczyk, M.A., Kuzniecky, R., Langner, S., Lenge, M., Leyden, K. M., Liu, M., Loi, R.Q., Martin, P., Mascalchi, M., Morita, M.E., Pariente, J.C., Rodriguez-Cruces, R., Rummel, C., Saavalainen, T., Semmelroch, M.K., Severino, M., Thomas, R.H., Tondelli, M., Tortora, D., Vaudano, A.E., Vivash, L., von Podewils, F., Wagner, J., Weber, B., Yao, Y., Yasuda, C.L., Zhang, G., Bargallo, N., Bender, B., Bernasconi, N., Bernasconi, A., Bernhardt, B.C., Blumcke, I., Carlson, C., Cavalleri, G. L., Cendes, F., Concha, L., Delanty, N., Depondt, C., Devinsky, O., Doherty, C.P., Focke, N.K., Gambardella, A., Guerrini, R., Hamandi, K., Jackson, G.D., Kalviainen, R., Kochunov, P., Kwan, P., Labate, A., McDonald, C.R., Meletti, S., O'Brien, T.J., Ourselin, S., Richardson, M.P., Striano, P., Thesen, T., Wiest, R., Zhang, J., Vezzani, A., Ryten, M., Thompson, P.M., Sisodiya, S.M., 2018. Structural brain abnormalities in the common epilepsies assessed in a worldwide ENIGMA study. Brain 141, 391–408.

Yogarajah, M., Powell, H.W., Parker, G.J., Alexander, D.C., Thompson, P.J., Symms, M. R., Boulby, P., Wheeler-Kingshott, C.A., Barker, G.J., Koepp, M.J., Duncan, J.S., 2008. Tractography of the parahippocampal gyrus and material specific memory impairment in unilateral temporal lobe epilepsy. Neuroimage 40, 1755–1764.

Zavaliangos-Petropulu, A., Nir, T.M., Thomopoulos, S.I., Reid, R.I., Bernstein, M.A., Borowski, B., Jack Jr., C.R., Weiner, M.W., Jahanshad, N., Thompson, P.M., 2019. Diffusion MRI Indices and Their Relation to Cognitive Impairment in Brain Aging: The Updated Multi-protocol Approach in ADNI3. Front. Neuroinform 13, 2.

Zhang, J., Cheng, W., Wang, Z., Zhang, Z., Lu, W., Lu, G., Feng, J., 2012. Pattern classification of large-scale functional brain networks: identification of informative neuroimaging markers for epilepsy. PLoS ONE 7, e36733.