A Framework for Brain Atlases: Lessons from Seizure Dynamics

Andrew Y. Revell[*,1,2,a], **Alexander B. Silva**[2,3,a], **T. Campbell Arnold**[2,3], **Joel M. Stein**[2,5], **Sandhitsu R. Das**[2,4], **Russell T. Shinohara**[6,7], **Danielle S. Bassett**[1,2,3,4,9,10,11], **Brian Litt**[2,3,4], **and Kathryn A. Davis**[1,2,4]

[1]Department of Neuroscience, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104 USA
[2]Center for Neuroengineering and Therapeutics, University of Pennsylvania, Philadelphia, PA 19104 USA
[3]Department of Bioengineering, School of Engineering and Applied Science, University of Pennsylvania, Philadelphia, PA 19104 USA
[4]Department of Neurology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104 USA
[5]Department of Radiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104 USA
[6]Department of Biostatistics, Epidemiology, and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104 USA
[7]Penn Statistics in Imaging and Visualization Endeavor, Perelman school of Medicine, University of Pennsylvania, PA 19104 USA
[8]Center for Biomedical Image Computing and Analytics, Perelman School of Medicine, University of Pennsylvania, PA 19104 USA
[9]Department of Electrical and Systems Engineering, School of Engineering and Applied Science, University of Pennsylvania, Philadelphia, PA 19104 USA
[10]Department of Physics and Astronomy, College of Arts and Sciences, University of Pennsylvania, Philadelphia, PA 19104 USA
[11]Santa Fe Institute, Santa Fe, NM 87501
[a]These authors contributed equally
[*]Corresponding author: andrew.revell@pennmedicine.upenn.edu

Understanding the relationship between the brain's structural anatomy and neural activity is essential in identifying the structural therapeutic targets linked to the functional changes seen in neurological diseases. An implicit challenge is that the varying maps of the brain, or atlases, used across the neuroscience literature to describe the different regions of the brain alters the hypotheses and predictions we make about the brain's function of those regions. Here we demonstrate how parcellation scale, shape, and anatomical coverage of these atlases impact network topology, structure-function correlation (SFC), and the hypotheses we make about epilepsy disease biology. Through the lens of our disease system, we propose a general framework to evaluate the validity of an atlas used in an experimental system. This framework aims to maximize the descriptive, explanatory, and predictive validity of these atlases. Broadly, our framework strives to augment neuroscience research utilizing the various atlases published over the last century.

Brain Atlas | Validity | Networks | Epilepsy | Structure-function

Introduction

How we define anatomical brain structures and relate those structures to the brain's function can either constrain or enhance our understanding of the biology of behavior and neurological diseases[1–4]. Discoveries by scientists like Carl Wernicke and Pierre Paul Broca who mapped specific brain regions to speech function, in addition to case studies from Phineas Gage and H.M. who lost specific brain regions with resultant changes in brain function and behavior, exemplify how brain structure and function are fundamentally linked[5–7]. Proper labeling of brain structures is paramount for effective communication amongst scientists about the variability between healthy individuals and about the regions involved in neurological disorders[8]. Yet no consensus has been reached on the most appropriate labeling and delineations of these regions, as manifest in the wide variety of brain maps or atlases defining neuroanatomical structures[9].

In common usage, an atlas refers to a "collection of maps"[10] that typically defines geo-political boundaries and may include coarse borders (continental), fine borders (city), and anything in between (country; Fig. 1a, left). Borders[11] are usually consistent across atlases of the world. In contrast, atlases of the brain are not consistent. Four separate atlases (Fig. 1a, right) may define the superior temporal gyrus differently. For example, over ninety percent of the *anterior* superior temporal gyrus in the Harvard-Oxford atlas[12] overlaps with the *posterior* superior temporal gyrus in the Hammersmith atlas[13]. Atlases may also differ in other ways, including the parcellation size, neuroanatomical coverage, and complexity of brain region shapes. For instance, the Yeo atlas[14] contains 7 or 17 parcels while the Schaefer atlases[15] may have between 100 and 1000 parcels. Complicating matters further, atlases can differ in their intended use. The MMP atlas[16] was intended for surface-based analyses[17], yet a volumetric version (without subcortical structures) was independently created and used in connectivity studies[18]. The plethora of available atlases poses a problem for reproducibility in the study of healthy and diseased populations and for metanalyses describing the involvement of different regions of the brain in various diseases. This has been termed the Atlas Concordance Problem[4].

In the present study, we perform a comprehensive evaluation of the available atlases in the neuroscience literature (Table 1) by examining the effect of varying features such as parcellation size, coverage, and shape (Fig. 1b) on structural connectivity (Fig. 1c) and structure-function correlation (SFC; Fig. 1d). Note the important distinction between the terms atlas, template, and stereotactic space[9] (see Fig. S1). In the context of our disease system, we propose a new framework outlining the validity of atlas used across experimental neuroscience systems. In our experimental design, we measure structural connectivity using high angular resolution diffusion imaging (HARDI) to capture the underlying anatomical connections between brain regions. We then measure neural activity using stereoelectroencephalography (SEEG) in epilepsy patients to capture real-time changes in seizure activity with finer temporal and spatial resolution than other functional neuroimaging modalities[20–22]. Finally, we utilize a total of 52 brain atlases freely available in common neuroimaging software to inves-

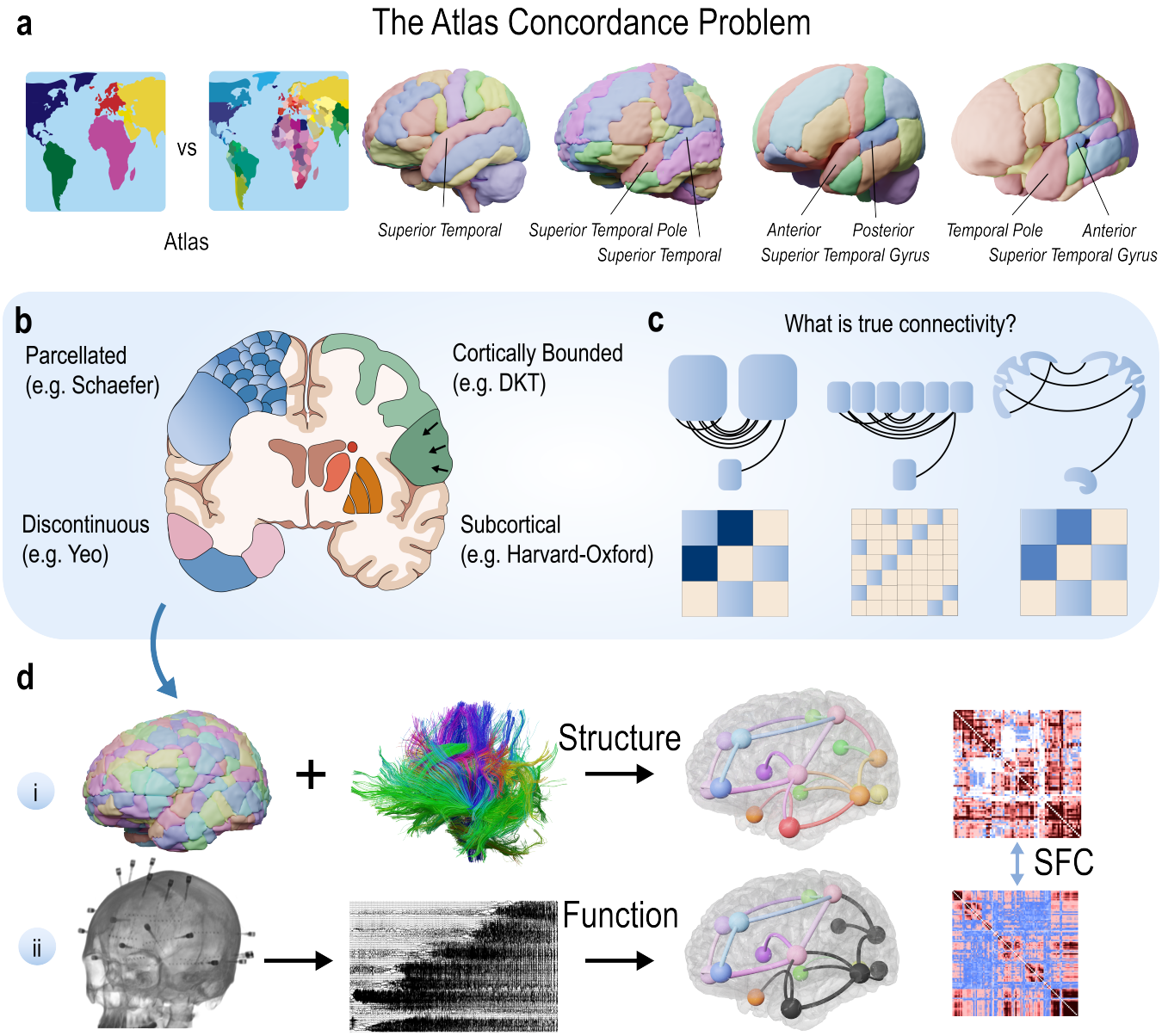

Fig. 1. The Atlas Concordance Problem and SFC. | **a**, In common usage, an atlas refers to a "collection of maps"[10] that defines geo-political boundaries. They may include coarse continental borders, fine state borders, or mesoscale country borders. Although borders[11] are usually consistent across atlases of the world, they are typically not consistent across atlases of the brain. Four separate atlases (left-to-right: CerebrA, AAL, Hammersmith, Harvard-Oxford) may define the superior temporal gyrus differently. The lack of consistency across these labels poses a problem for reproducibility in cognitive, systems, developmental, and clinical studies, as well as metanalyses describing the involvement of different regions of the brain of various diseases[4]. This challenge has been previously referred to as the Atlas Concordance Problem. **b**, Atlases can have varying features (see also Table 1). **c**, The varying definitions of anatomical areas decreases confidence that all current connectivity studies reflect some fundamentally "true" architecture. **d**, When combined with white matter tracts reconstructed from diffusion MRI, atlases can be used to measure how different regions of the brain are structurally connected (i). Similarly, intracranial EEG (iEEG) implants can record neural activity to measure how different regions of the brain are functionally connected (ii). The statistical similarity between structural and functional connectivity measurements can be calculated (e.g., structure-function correlation; SFC), and such estimates have recently been used to better understand the pathophysiology of disease.

tigate hypotheses about the structure-function relationship in epilepsy patients. We found parcellation scale affects the measurement of resting-state SFC (rsSFC) and the change in SFC (ΔSFC) at seizure onset, potentially altering conclusions about how seizures harness the underlying structural scaffold in the brain supported in prior research[23–26].

Atlas [regions]	Sources	3D Render	Description	Variations
AAL [116;120;166]	1-7 SPM	S	Structural atlas. Manual identification using a defined labeling protocol on single subject template (Collin-27). Three versions. Version 2: updated boundaries. Version 3: further parcellations. Successor to Talairach.	AAL: AAL1, AAL2, AAL3, AAL600, AAL-JHU AAL1 AAL2 AAL3 AAL600 Removed (dark blue) Smaller (light blue) Added (red-yellow) AAL-JHU (JHU labels blue)
AICHA [384]	8	F	Functional atlas based on rsfMRI; 281 subjects. Each ROI has (1) homogeneity in its functional activity (2) a homotopic contralateral counterpart with which it has maximal connectivity.	
Brainnetome [246]	9-10 DSIstudio	S	Connectivity-based parcellation. Based on idea that clustered regions of a brain region should share similar connectivity profiles; 40 subjects from HCP dataset. 210 cortical; 36 subcortical.	
Brodmann [48]	11-13 MRIcron	S	Developed by independent group at Washington University in St. Louis. Published with MRIcron software. Warned by developer to be used with caution - not validated, nor based on multiple individuals.	
CerebrA [102]	14	S	Structural atlas. Non-linear registration of cortical and subcortical labelling from Mindboggle-101 dataset (see DKT below) to the symmetric MNI-ICBM2009c template, followed by manual editing.	Craddock: N parcellations N=200 N=400 1.7 cm 1.0 cm pea
Craddock [N]	15-17	F	Functional atlas; rsfMRI; 41 subjects. ROIs are spatially clustered into regions of homogeneous functional connectivity. May be N regions. 200/400 regions publicly available. 4x4x4 mm^3 resolution fMRI. Resliced.	
DKT [111]	18-23 Freesurfer	S	DKT is a labelling *protocol*. Used on Mindboggle-101 dataset (101 brains). Probabilistic atlas created using joint fusion algorithm. Surface version in Freesurfer. Volumetric version uses 20 brain subset. Noncortical regions: Neuromorphometrics BrainCOLOR atlas.	Harvard-Oxford: Cortical/subcortical only, combined, symmetric, nonsymmeric Symmetric Nonsymmetric Subcortical Combined Cortical + Subcortical
Gordon-Petersen [333]	24-25	F	Identification of abrupt transitions in resting-state functional connectivity to identify parcellations. Based on rsFMRI. 108 subjects. Intended for surface-based analyses.	
Hammersmith [83]	26-28	S	Manually identified 83 structures using defined labelling protocol; 30 subjects. Maximum probability map. First version in 2003 with 49 structures. Named after London hospital, Hammersmith. Hammers is author.	
Harvard-Oxford [48 + 21]	29-30 FSL	S	Manual segmentation using defined labelling protocol; 37 subjects. Cortical and subcortical atlases provided separately. Left and right structures have same labels (symmetry). Must preprocess.	
JHU [48; 20]	31-33 FSL	S	White matter atlas. Two versions. (1) Labels: Hand segmentation average of diffusion MRI; 81 subjects. (2) Tracts: probabilistic identification from deterministic tractography; 28 subjects.	JHU: Labels, tracts Labels Tracts
Julich [121]	34-35 FSL	S	Cytoarchitecture atlas. Successor to Brodmann. Average of 10-subject post-mortem cyto- and myelo-architectonic segmentations. Update to the Eickhoff SPM Anatomy Toolbox v1.5. Whole brain is not covered.	
MMP [380]	36-38 DSIstudio	M	Multi-modal parcellation: (1) Architecture - T1w/T2w myelin maps + cortical thickness, (2) function - task-fMRI, (3) connectivity, (4) topography. 210 subjects. Cortical ONLY. Originally intended for surface analysis. Volumetric version independently created and used.	Random: N parcellations, cortical, whole-brain, subparcellated N=10 N=30 N=100 N=1,000 N=10,000 lemon 5 cm grape 2 cm pea 1 cm
Random [N]	39-40	V	Brain is randomly parcellated into N regions. Variations used in studies include cortical and whole-brain. Other atlases (e.g. AAL) and their regions may be further randomly divided, or subparcellated.	
MNI Structural [9]	41 FSL	S	9 regions, including lobar and some subcortical regions. Hand segmented 50 subjects. Transformed into MNI152 space, averaged, probability maps produced. 25% max probability is shown.	Schaefer: 100 to 1,000 parcellations (by 100), named to Yeo 7 and 17 N=100 N=500 N=1,000
Schaefer [100-1000]	42-43 Github	F	Based on rsfMRI. Clusters found with gradient-weighted Markov Random Field model. 1489 subjects. Cortical only. Spatial resolutions provided: 100 - 1000 parcellations (by 100). Well documented.	
Talairach [1105]	44-48 FSL	S	Conversion of original Talairach labeling. Digitized version of the original (coarsely sliced) Talairach atlas and registration to MNI 152 space. Atlas provided in FSL.	Yeo: 7/17 parcellations; Cortically bounded or liberal Cortically bounded liberal discontinuous
Yeo [7; 17]	49-50 Freesurfer	F	1000 subjects; rsfMRI. Clustered cortical regions by pattern of functional connectivity. Results in non-spatially continuous clusters. 7 and 17 clusters based on stability of clustering algorithm.	
Region-specific	41-54 FSL	V	Atlases created for specific regions, usually high quality + high degree of accuracy (e.g. post-mortem histological verification). Examples: Thalamus nuclei, hippocampus, and other specific structures.	Thalamus, Hippocampus, Cerebellum Cerebellum
Population-specific	55-56	V	Atlases created from a specific population (e.g. elderly, pediatric, non-human). Disease-specific defines regions specific for disease (e.g. MS lesion probabilistic locations).	Pediatric, Elderly, Disease specific Neonatal M-CRIB (Melbourne)

Table 1. Atlases. | Refer to Table S1 for atlas sources. NIfTI files converted to STL files with Slicer and 3D rendered in Blender. **S**: Structurally defined atlas; **F**: Functionally defined atlas; **M**: Multi-modally defined atlas; **V**: A variably defined atlas that may be structural, functional, multi-modal; **rsFMRI**: resting-state fMRI; **ROI**: region of interest; **HCP**: Human connectome project dataset [19]; **DKT**: Desikan-Killiany-Tourville protocol [1]; **MS**: multiple sclerosis.

Through the lens of our disease system, we conclude with a new framework for evaluating atlases by expanding historical foundations for assessing the validity and effectiveness of animal models [27], network models [28], and psychometric tests [29]. A one-size-fits-all approach may not nor should exist [30]. Instead, we hope to critically evaluate an atlas by maximizing its (1) descriptive, (2) explanatory, and (3) predictive validity [28] in relation to the experimental system at hand. In epilepsy specifically, we aim to select an atlas that resembles the system in which we work (**descriptive validity**). Importantly,

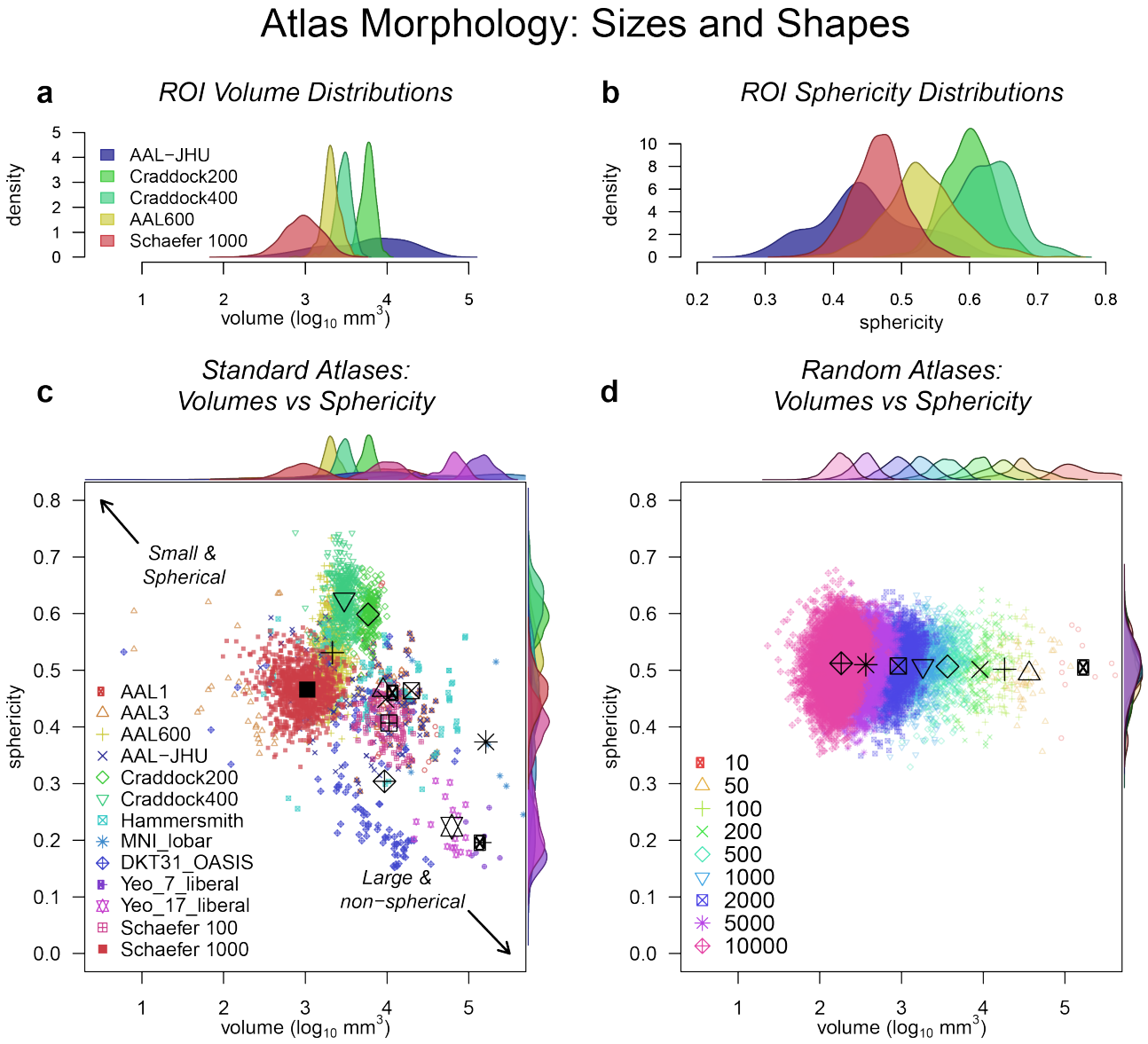

Fig. 2. Atlas morphology: sizes and shapes. | **a**, Volume distribution of atlas parcellations or region of interests (ROIs) demonstrating the diversity of parcellation sizes. Some atlases have wide distributions, while others have low variability in parcellation sizes. **b**, Parcellation sphericity distributions illustrating how the shapes of different parcellations may not be uniform. Colors denoting atlases are the same in **a**. **c**, Volumes versus sphericity showing how some atlas parcellations may be small and spherical, while others may be large and non-spherical. This illustrates the non-uniformity in atlas parcellations. We hypothesize that this variability contributes to altered network structure and measurement of SFC. d, Volumes and sphericity of random atlases showing the uniformity of sphericity with changing volumes. This allows us to study the effect of parcellation scale on network characteristics and SFC without the confound of shape effects. Numbers in legend represent the number of parcellations for each random atlas. Remaining atlases are in Fig. S2. See Table S1 for atlas descriptions.

it should include coverage of subcortical structures typically involved in epilepsy networks with a parcellation scale that is not too coarse nor fine to model connectivity at the appropriate scale (given the resolution limits of HARDI and SEEG). Next, we want to select an optimal atlas that can be used for hypothesis testing (**explanatory validity**). It should include the capability to test how functional changes seen in epilepsy are related to the underlying structural connectivity. Explanatory validity thus requires assessment of both the atlas features (a form of descriptive validity) and its ability to test for causal relationships. Finally, we strive to maximize the predictive capability of an atlas (**predictive validity**). We aim to optimally predict functional changes seen in epilepsy using noninvasive structural neuroimaging, lessening the need for costly and invasive implantations in epilepsy patients. Later, we show some atlases at a particular scale are not able to pre-

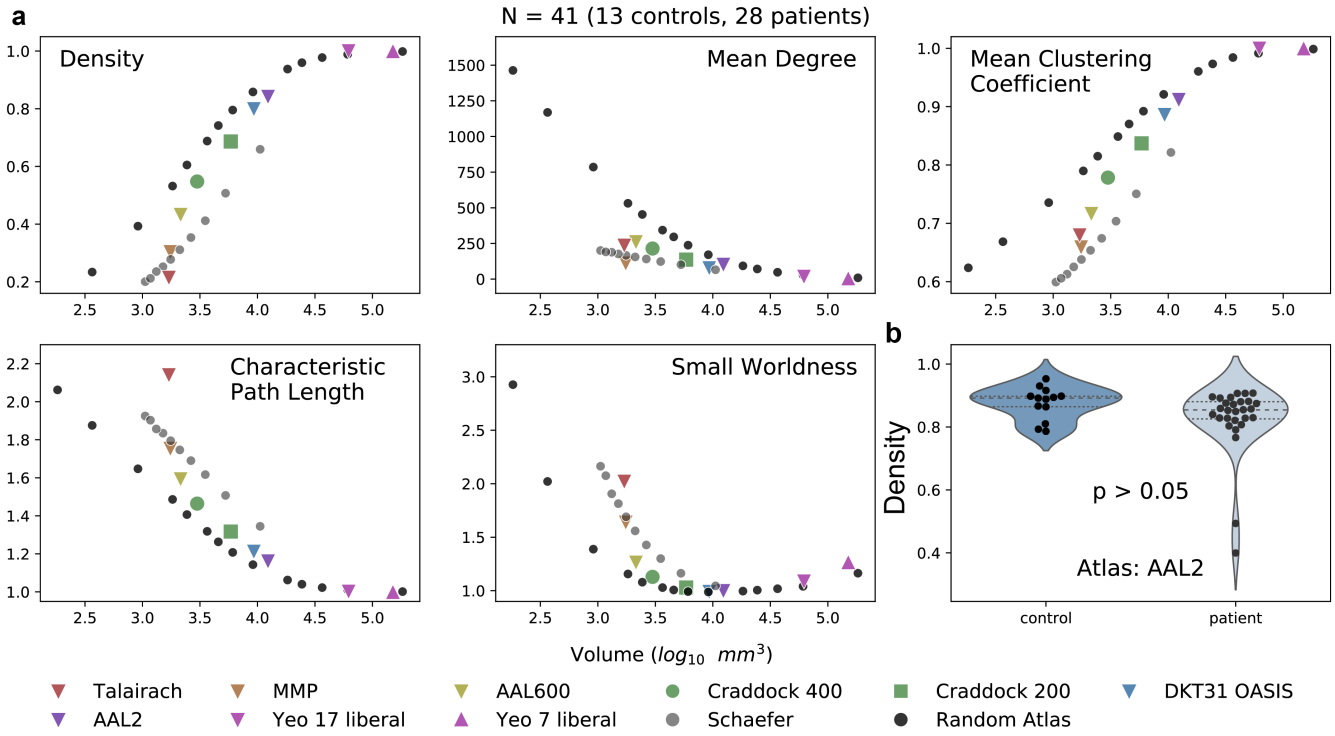

Fig. 3. Network differences between atlases. | **a**, Density, mean degree, mean clustering coefficient, characteristic path length, and small worldness were calculated for all atlases. A subset of atlases is shown. Remaining atlases studied are shown in Fig. S3. The average parcellation volume was calculated for each atlas and the corresponding network measure was graphed as the mean of all subjects (N=41; 13 controls, 28 patients). Because many basic network measures change as a function of parcellation volume, we hypothesized that SFC would also change based on parcellation volume. **b**, Controls and patients were not significantly different in density for the AAL2 atlas (Mann-Whitney U test) illustrating global structural network measures are similar between cohorts. Specific connectivity differences between cohorts were not explored (e.g. to explore if connections from the hippocampus to the cingulate gyrus are changed in temporal lobe epilepsy) and out of the scope of this manuscript. Controls and patients were separated and shown in Fig. S4. See Table S1 for atlas descriptions.

dict structure-function changes with seizure onset. With this framework, the present study demonstrates the set of atlases with specific features such as parcellation size, shape, and coverage that meet our goal of predicting functional changes seen in epilepsy. Not all atlases are valid for a specific study. Our generalized framework provides a valuable resource for others to make an educated decision in regards to atlas choice when designing their study.

Results

Clinical Data. Forty-one individuals (mean age 34 ± 11; 16 female) underwent High Angular Resolution Diffusion Imaging (HARDI), composed of thirteen controls (mean age 35 ± 13; 6 female) and twenty-eight drug-resistant epilepsy patients (mean age 34 ± 11; 12 female) evaluated for surgical treatment. Of the twenty-eight patients, twenty-four were implanted with stereoelectroencephalography (SEEG) and four with electrocorticography (ECoG). Ten SEEG patients (mean age 34 ± 8; 4 female) had clinical seizure annotations, and the first seizure from each patient (mean duration 81s) without artifacts was selected for SFC analyses. Patient and control demographics are included in Table S2.

Atlas Morphology: Sizes and Shapes. We hypothesized that atlas morphological properties, including size and shape, affect SFC. To test this hypothesis, we first quantified the distributions of parcellation sizes and shapes in various atlases (Fig 2). Some atlas parcellations have narrow volume distributions (Fig 2a, e.g. Craddock 200 and 400 atlases), while others have wider parcellation volume distributions (e.g. Schaefer 1000). Several atlases that have larger parcellation volumes may have lower sphericity values (Fig 2b). These results exemplify the diversity of atlas parcellation morphology. Fig 2c shows a comparison of individual parcellation volumes and sphericities. The remaining atlases are shown in Fig. S2. In contrast to standard atlases, random atlases have constant sphericity with respect to a change in volume size. Although random atlases may not represent true anatomical or functional boundaries, the benefit is that the shape of a parcellation is uniformly biased regardless of parcellation size; random atlases allow us to study how parcellation scale affects network structure and SFC while keeping the effect of shape constant. They also allow us to explore if accurate and precise anatomical boundaries are crucial for our experimental system.

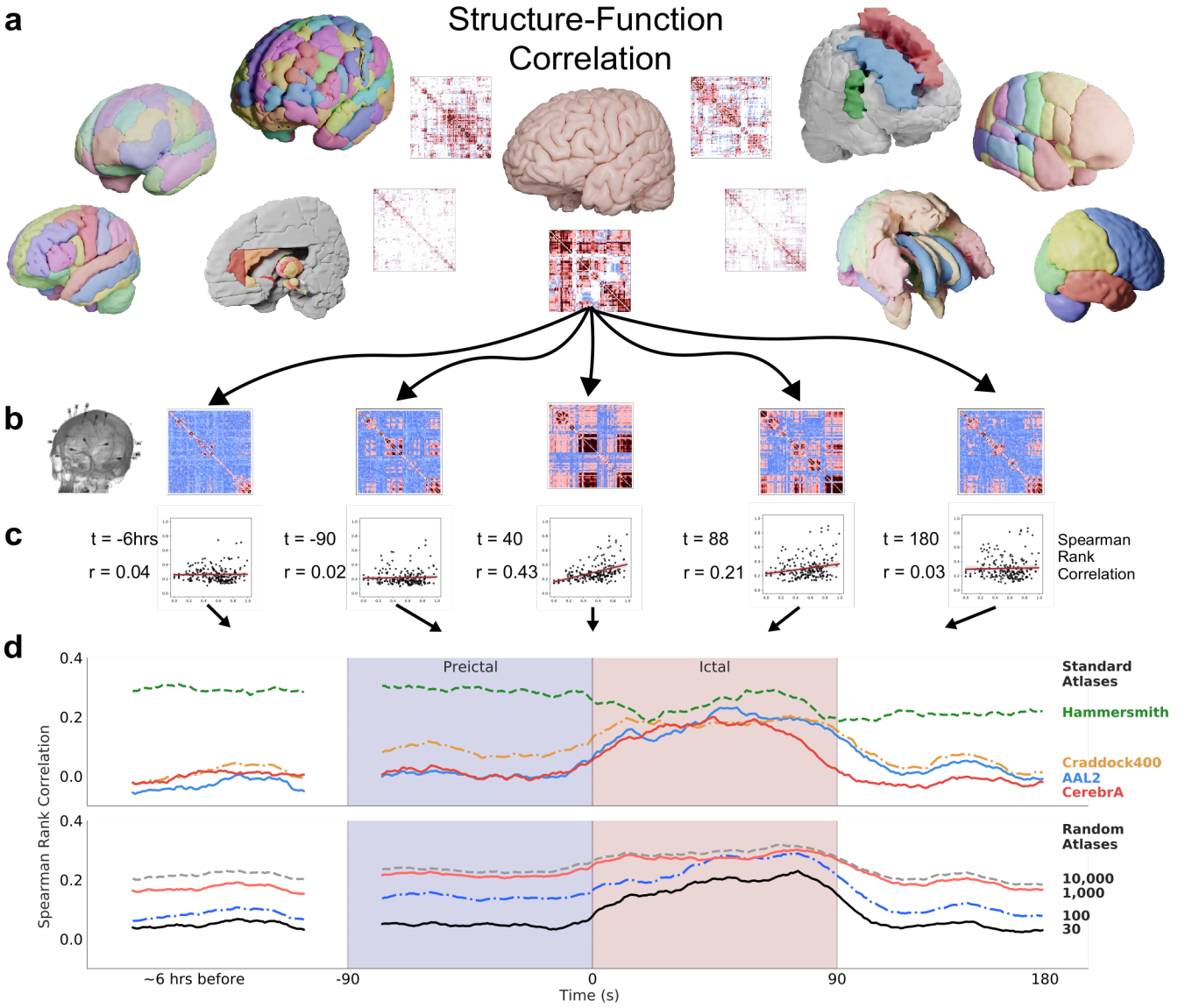

Fig. 4. Structure-Function correlation in a single patient. | **a**, Example atlases and structural adjacency matrices. **b**, Functional adjacency matrices are produced from the analysis of SEEG recordings during the interictal, preictal, ictal, and postictal periods. Broadband cross correlation matrices are shown for a single patient RID0278 at 6 hours before seizure onset, 90 seconds before seizure onset, 40s after seizure onset (t = 40), 88 seconds after seizure onset (seizure duration = 89 seconds), and 180 seconds after seizure onset (or 91 seconds after seizure termination). **c**, Each functional adjacency matrix is correlated to a structural adjacency matrix of a given atlas. A plot of the structural edge weights and corresponding functional edge weights is shown for the example time points of **b**. Spearman Rank Correlation is measured between all time points and all atlases for each patient. **d**, SFC is graphed at each time point for four example standard atlases (Hammersmith, Craddock400, AAL2, and CerebrA), and four example random whole-brain atlases (30, 100, 1000, and 10000 parcellations). SFC increases during seizure state for some standard atlases – Craddock 400, AAL2, and CerebrA atlases. This result follows previous SFC publications with ECoG[23,24]. However, SFC does not increase for the Hammersmith atlas. These findings highlight inference from one type of atlas may suggest that seizure activity is not correlated to brain structure, contradicting previous studies. Similarly, SFC increases for a subset of random whole-brain atlases. See Table S1 for atlas descriptions.

Anatomical definitions affect network topology. Although the morphology of atlas parcellations is diverse, we aimed to investigate how these morphological characteristics affect network topology, particularly how parcellation scale affects network structure Fig. 3. Networks are the basis upon which we compute SFC, and not necessarily morphological characteristics, therefore, we measured how network density, mean degree, characteristic path length, mean clustering coefficient, and small worldness change as a function of parcellation scale (Fig. 3a). We found that the change in these network mea-

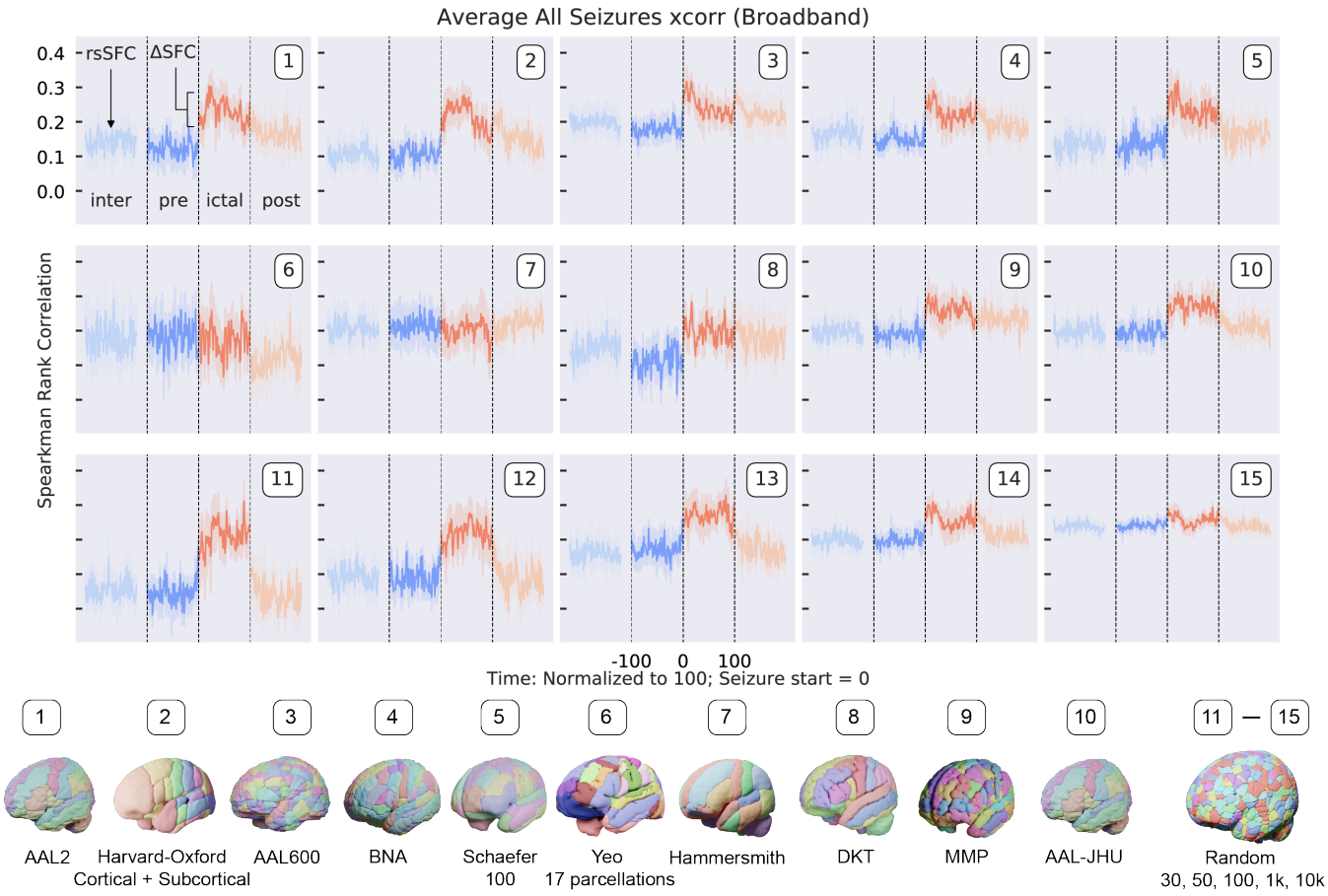

Fig. 5. Structure-Function Correlation in multiple patients. | SFC for ten standard atlases and five random atlases using SEEG broadband cross-correlation metrics averaged across all patients with clinically annotated seizures (N = 10). Resting state SFC (rsSFC) is the SFC during the interictal period. We observe that rsSFC increases with decreasing parcellation volume (see also Fig. 6a). The AAL2 atlas shows a statistically significant increase in SFC from preictal to ictal periods ($p = 0.02$ by Wilcoxon signed rank test after Bonferroni correction for 52 tests). The change from preictal to ictal SFC is ΔSFC. This finding supports the hypothesis that seizure activity harnesses the underlying structural connectivity of the brain. SFC was similarly calculated for random whole-brain atlases. These findings may be concerning given that the *inherent* structure-function relationship in the brain is not necessarily changing at resting state, but its measurement is greatly affected by atlas choice alone. These results highlight the crucial need for critically evaluating the appropriate atlas to understand SFC across the neuroscience literature, especially in an SEEG setting given the rise of SEEG implantations30. **xcorr**: cross correlation. See Table S1 for atlas descriptions.

sures are congruent between standard and random atlases and previous studies[31]. For example, density and mean clustering coefficient increase as a function of increasing average parcellation volume for both the standard and random atlases, while characteristic path length and small worldness decrease. We also show that mean density, a global network measure, is similar between our control (N=13) and patient (N=28) cohorts (Fig. 3b). As a result of these findings, we hypothesized that SFC would also change based on parcellation volume.

Anatomical definitions affect SFC. Fig. 4 illustrates an overview of how SFC is calculated. Structure is measured with high angular resolution diffusion imaging (HARDI) and function is measured with SEEG electrode contacts. Structural adjacency matrices are generated based on the atlas chosen (Fig. 4a) and functional adjacency matrices are generated based on broadband (1 – 128 Hz) cross-correlation of neural activity between the electrode contacts (Fig. 4b). The adjacency matrices shown are example data from a single patient, RID0278. Functional adjacency matrices for RID0278 are shown for 6 hours before seizure onset, 90 seconds before seizure onset (t = -90), 40 seconds after seizure onset (t = 40), 88 seconds after seizure onset (seizure duration = 89 seconds), and 180 seconds after seizure onset (91 seconds after seizure termination). Each functional adjacency matrix was correlated to each structural adjacency matrix, yielding a SFC at each time point (Fig. 4c). Each point represents the normalized structural edge weight between two brain regions and their corresponding functional connectivity edge weight in broadband cross-correlation. A line of best fit is shown, and r values represent Spearman rank correlation for that time point. SFC was graphed for all time points during the interictal, preictal,

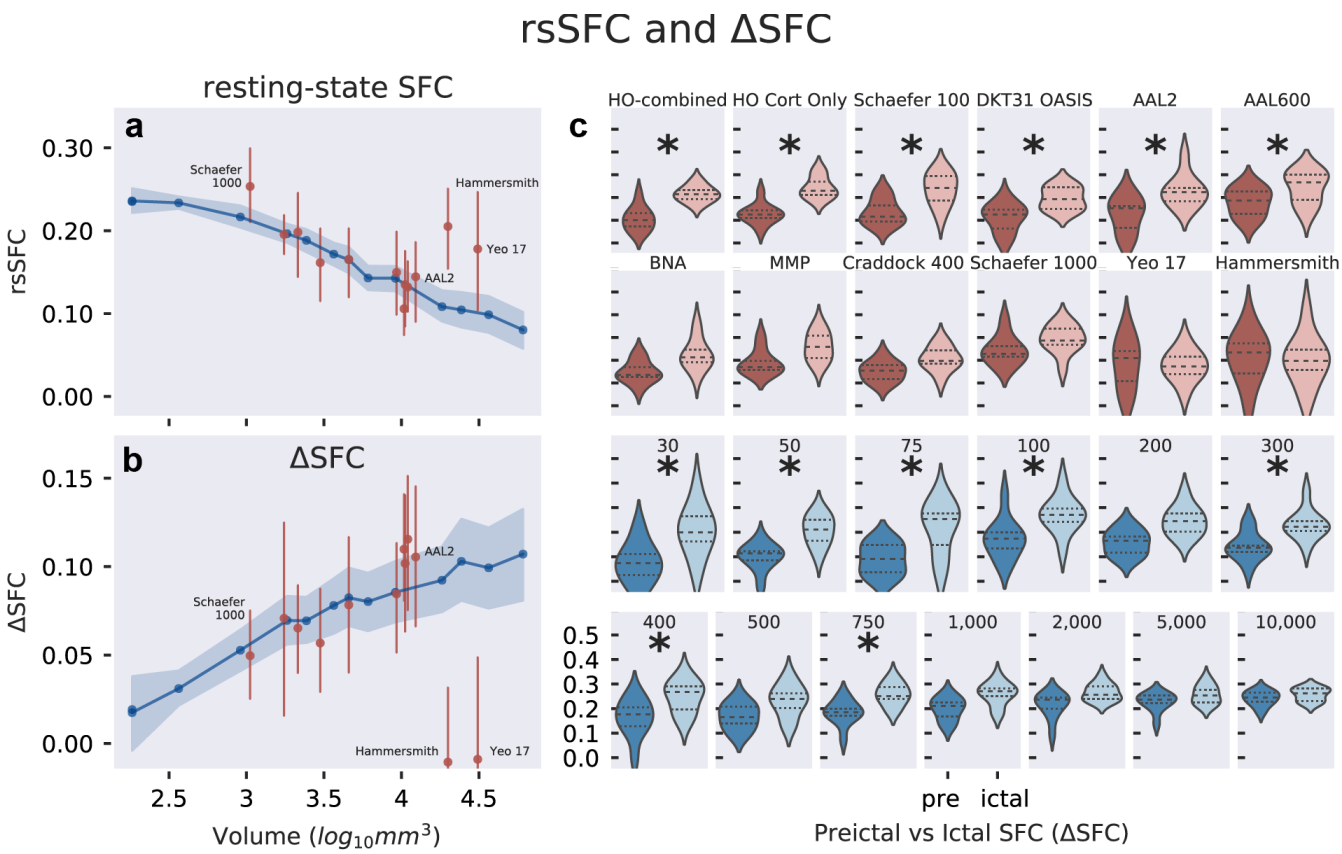

Fig. 6. A trade-off between resting-state SFC (rsSFC) and the change in SFC (ΔSFC) with neuroanatomical scale. | **a**, rsSFC increases at smaller parcellation scales. Random atlases are shown in blue and standard atlases are shown in red. Bands represent 95% confidence intervals. **b**, ΔSFC decreases at smaller parcellation scales. Broadly, ΔSFC may be interpreted as the change in SFC with respect to disease (e.g. seizure, schizophrenia, major depressive disorder) and non-disease states, and this change has been used to characterize and make inferences on many neurological diseases. These results exemplify that either too coarse or too fine parcellations may not adequately capture the underlying SFC of the brain or its dynamics with relation to neurological disease. **c**, A subset of atlases capture the dynamical change in SFC. The Harvard-Oxford (HO) and DKT atlases show a significantly different SFC between preictal and ictal periods ($p < 0.05$ by Wilcoxon signed rank test after Bonferroni correction for 52 tests) while the Brainnetome and MMP atlases do not ($p > 0.05$). Larger parcellation volumes (e.g. N = 30) increase in ΔSFC from smaller parcellation volumes (e.g. N = 1,000), indicating that larger parcellation volumes adequately capture ΔSFC. However, parcellation volume is not the only factor in adequately capturing ΔSFC. The Hammersmith atlas with large parcellation volumes and larger electrode coverage is not able to capture a significant ΔSFC ($p > 0.05$). Asterisks represent atlases with statistically significant differences in SFC between ictal and preictal periods after Bonferroni Correction. See Table S1 for atlas descriptions.

ictal, and postictal periods for this patient in Fig. 4d.

Four example standard and random atlases are graphed. We show that SFC increases during the ictal state for many atlases (CerebrA, AAL2, Craddock 400), but not all atlases (Hammersmith). The increase in SFC during seizures follows previous SFC publications with ECoG[23,24]. Similarly, SFC increases for a subset of random whole-brain atlases. The random whole-brain atlases were created to change parcellation scale while preserving shape, therefore, these data support that SFC is affected by parcellation scale. However, parcellation scale is not the only feature affecting SFC – the Hammersmith and AAL2 atlases have similar parcellation scales yet diverging neuroanatomical properties and SFC dynamics. These findings highlight inference from one type of atlas may suggest that seizure activity is not correlated to brain structure, contradicting previous studies[23]. Broadly, structural network studies and conclusions may be affected by the atlas chosen, and thus care must be taken when interpreting the structure-function relationship of the brain with respect to neuroanatomical definitions.

Structure-Function Correlation at a Population Level. Fig. 5 shows SFC for ten standard atlases and five random atlases using SEEG broadband cross-correlation metrics averaged across all patients with clinically annotated seizures (N = 10). Functional connectivity measurements were also calculated for coherence, zero time-lag Pearson, and Spearman rank correlations across multiple frequency bands. They are included in the freely available, curated, and opensource dataset for all readers of this manuscript (see methodology section for

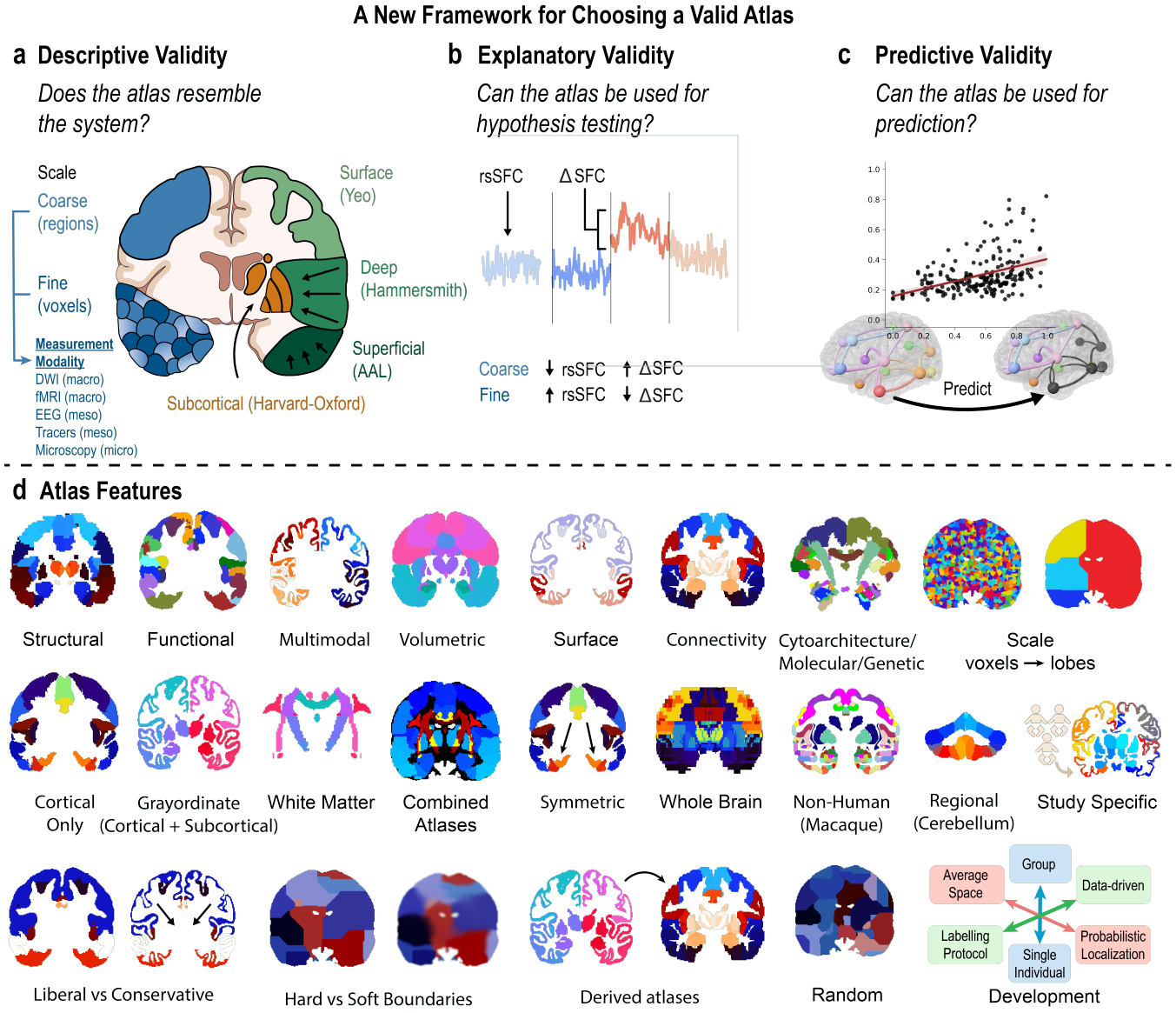

Fig. 7. Framework for selecting a brain atlas. | **a**, Descriptive validity of an atlas addresses the "face value" of an atlas (it resembles the experimental system). In epilepsy, an atlas should cover subcortical structures involved in seizure generation. It should have a parcellation scale not too coarse nor fine to model connectivity at the appropriate scale given the resolution limits of HARDI and SEEG. **b**, Explanatory validity of an atlas addresses whether an atlas can be used for testing causal relationships. In epilepsy, it includes the capability to test how functional changes seen in epilepsy are related to structural connectivity. **c**, Predictive validity addresses the predictive capability of an atlas. In epilepsy, we want to be able to predict functional epileptic changes using noninvasive neuroimaging, lessening the need for costly and invasive implantations. **d**, Non-mutually exclusive atlas features related to descriptive validity. Atlases may be derived from structural, functional, or multimodal datasets; parcellations can be random. Atlases may be volumetric or surface-based, made from connectivity data (structural/functional), derived from microscopic datasets such as cytoarchitecture, molecular, and genetic. Parcellations may range from voxels to entire lobes, have different brain coverages (cortical, "grayordinate", white matter, or whole brain), may be combined, and have symmetric labeling. Atlases may be non-human, incorporate highly detailed maps of specific regions, and be made from the study participants. Parcellations may include liberal/conservative and hard/soft boundaries. Atlases can be derived and further parcellated from other atlases. Finally, atlases may be developed through different methodologies with further details in the text.

links); however, they were not used in directed hypothesis testing about specific frequency bands nor about other functional measurements in the present study. The AAL2 atlas shows a statistically significant increase in SFC from preictal to ictal periods ($p < 0.05$ by Wilcoxon signed rank test after Bonferroni correction for 52 tests). The change from

preictal to ictal SFC is denoted ΔSFC. This finding supports the hypothesis that seizure activity harnesses the underlying structural connectivity of the brain [23,24]. SFC was similarly calculated for random whole-brain atlases. A notable finding is that during the interictal period, resting state SFC (rsSFC) increases at larger number of parcellations (i.e. smaller parcellation volumes). We show that rsSFC is observably affected by parcellation scale when inspecting the random atlases in Fig. 5 (bottom row).

These findings may be concerning given that the *inherent* structure-function relationship in the brain is not necessarily changing at resting state, but its measurement is greatly affected by atlas choice alone. These results highlight the crucial need for critically evaluating an appropriate atlas, particularly by maximizing explanatory validity. We aim to maximize our ability to test hypothesises and draw conclusions about structure-function relationship within the brain at seizure onset. This is especially important in an SEEG setting given the rise of SEEG implantations worldwide; it allows for sampling of cortical and subcortical structures across both hemispheres with reduced morbidity and higher tolerance for patients [32].

rsSFC vs ΔSFC. Resting state SFC (rsSFC) and the change in SFC (ΔSFC) from preictal to ictal periods are differentially affected by parcellation scale (Fig. 6). Fig. 6a shows how rsSFC increases at smaller parcellation scales. In contrast, Fig. 6b shows how ΔSFC decreases at smaller parcellation scales. Broadly, ΔSFC may be interpreted as the change in SFC with respect to disease (e.g. seizure, schizophrenia, major depressive disorder) and non-disease states. This change metric has been used to characterize and make inferences in many neurological disorders [33,34]. These results exemplify that either overly coarse or fine parcellations may not adequately capture the underlying SFC of the brain or its dynamics with relation to neurological disease (low explanatory validity in our framework below).

A subset of atlases can capture the dynamical change in SFC (Fig. 6c). For example, the Harvard-Oxford (HO) and DKT atlases show a significantly different SFC between preictal and ictal periods ($p < 0.05$ by Wilcoxon signed rank test after Bonferroni correction for 52 tests) while the Brainnetome and MMP atlases do not ($p > 0.05$). Larger parcellation volumes (e.g. N = 30) result in an increase in ΔSFC compared to smaller parcellation volumes (e.g. N = 1,000), indicating that larger parcellation volumes adequately capture ΔSFC. However, parcellation volume is not the only factor in optimally capturing ΔSFC. The Hammersmith atlas with large parcellation volumes and larger electrode coverage (Supplementary Fig. 4) is not able to capture a significant ΔSFC ($p > 0.05$).

Discussion

In this study, we performed a comprehensive evaluation of the available structural, functional, random, and multi-modal atlases in the neuroscience literature (Table 1). We detailed morphological and network differences between these atlases and showed the effect of varying neuroanatomical definitions on the measurement of structure-function correlation (SFC) in epilepsy patients. We showed how the various atlases may alter conclusions about seizure dynamics. This work has wide implications for neuroscience labs utilizing such atlases because some atlases produce different results and may alter predictions and conclusions we draw about the brain's function. Based on our study, we propose a general framework below for evaluating and selecting atlases (Fig. 7a-c) to direct future neuroscience work.

A New Framework for Brain Atlases. Various publications have highlighted the atlas concordance problem [2–4,9], curated several atlases in freely accessible databases [35,36], and have made arguments for why specific atlas features (Fig. 7d) may be valid or superior in certain situations [17,30,37–41]. Clearly, there have been great efforts to publish accurate and precise parcellations both in individuals and across populations as seen with an exponential rise in atlas-related publications over the last three decades (Fig. S6). However, none have found a general solution to the underlying problem: Does atlas choice matter? If so, how much? And are there atlas features important in certain situations or experimental designs? An argument can be made that in some cases, atlas choice may not matter to a great extent. For example, many atlases show similar results in Fig. 6. Interestingly random whole-brain atlases, which do not follow accurate anatomical or functional boundaries, show SFC changes at seizure onset in concordance with other common atlases used across neuroscience studies. We provide a general framework that allows us to determine if an atlas is valid, avoid testing large numbers of atlases at one extreme, and gives credence to the current standard of publishing results in the main text using a single atlas and, if deemed necessary, provide supplementary results using a different atlas (preferably one with different features). A one-size-fits-all standard parcellation may not nor should exist [30].

Our framework evaluates the validity of an atlas to maximize its (1) descriptive, (2) explanatory, and (3) predictive validity [28] in relation to the experimental system. This framework is borrowed from the logic of assessing network models, and historically, animal models [27,42] and psychometric tests [29,43], where assessment of these models with standard statistical model-selection methods is particularly challenging.

Descriptive validity of an atlas refers to an atlas that appropriately resembles the system in which we work. In other words, it has "face value" [27]. This includes atlas features (Fig. 7d) relevant to the study, for example, the inclusion of subcortical structures relevant to epilepsy. Without the inclusion of relevant features, an atlas may not allow for hypothesis testing or determination of causality (explanatory validity below). Importantly, descriptive validity of an atlas also relates to the modality scale we use to measure the brain – for example, DWI and fMRI at the macroscale [44], iEEG and tracers at the meso scale [45], and microscopy at the microscale [46]. It is important to select a parcellation scale that resembles the measurement modality resolution (Fig. 7a). When correlating DWI with iEEG in our study at larger parcellations, we lose our ability to discern precise anatomical locations that are structurally and functionally related. Furthermore, resting state structure-function relationship increases at larger parcellations even though the *inherent* structure-function relationship should remain constant (Fig. 5 and Fig. 6a). At the other end of the scale, we lose our ability to discriminate the differences in structure-function relationship at seizure onset at smaller parcellations (tending to the size of voxels). At voxel scales, other notable limitations include problems with

multiple comparisons, computational costs, (near) collinearity, and the introduction of noise with inaccurate alignment of individual subjects' data. Recommendations for performing voxel-level analyses versus larger node-based approaches are discussed in Bijsterbosh et al. 2017 [47].

Two additional atlas features highlighted here include (1) surface versus volumetric based atlases and (2) atlas development. Surface based registration may improve accuracy over traditional volumetric based approaches [17], however, a limitation is that a large proportion of brain activity involves communication between cortical and subcortical regions and thus a surface-based approach is unlikely to provide a complete understanding of the brain [47]. Although the brain including the cortex, with six cytoarchitecturally defined layers, is fundamentally not a surface, we showed how surface based atlas features may not be as vital for consideration as features such as scale (Fig. 6c; Yeo, and DKT being surface-based and AAL and Harvard-Oxford being volumetric-based). A combination of surface cortical and sub-cortical gray matter regions, or "grayordinate" [19] atlases may be appropriate in some cases. Thus the atlas chosen relies on consideration of the experimental system. Finally atlases may have been developed through three non-mutually exclusive axes (Fig. 7d, bottom right): (1) using a single representative individual (Talairach) to a group of individuals (e.g. Hammersmith) [3]; (2) using a human labeling protocol (AAL) to a data-driven approach (e.g. Yeo); and (3) using a standard space representation such as in MNI coordinate space (Harvard-Oxford) to using a probabilistic mapping of the study participants (DKT in Freesurfer). The last approach where atlas labels are manually annotated and used as training classifiers to label the study participant brains is notably time consuming and is limited in use across studies [38,48].

Explanatory validity of an atlas requires an assessment of both an atlas' architecture (a form of descriptive validity) and its ability to test for causal relationships [28]. In epilepsy, it includes the capability to test hypotheses on how functional changes are related to the underlying structural connectivity, if at all. Statistical testing, such as tests to determine if specific brain regions are significantly different from each other in controls and patients, and subsequent conclusions drawn from the use of an atlas is the focus of explanatory validity. With explanatory validity, the biases introduced into our results from using an atlas must also be acknowledged, which can alter our conclusions about neurobiology and pathophysiology. For example, some structural atlases have different anatomical labeling protocols (DKT, AAL, Hammersmith) which may introduce biases resulting from how large, small, or the exact anatomical landmarks were used to create such atlases. Data-driven atlases [14,16,47], namely those created through functionally related brain regions, may also introduce biases based on the measurement modality or nodal definitions used (clustering, decomposition, gradient-based methods) and alter Type I or Type II error rates. In our study, we investigate whether seizures spread through the underlying connectome of the brain at the macro-scale level and if structural connectivity can be used to predict seizure spread. If seizures spread along the human connectome, but an atlas with >1,000 parcellations shows no change in SFC at seizure onset, we may introduce a Type II statistical error.

Predictive validity of an atlas indicates the ability of a certain measure to predict some other criterion measure [49]. For example, it can be incorporated into an analysis pipeline to predict a change in response to a perturbation, such as a drug, electrical or chemical stimulation, or a dynamical disease state. In our study, the perturbation is the change in brain state at seizure onset. Predicting functional changes in epilepsy using noninvasive structural neuroimaging is particularly useful clinically and will lessen the need for costly and invasive implantations in patients. An atlas that adequately captures ΔSFC with seizure onset will allow us to form network models to predict seizure related activity in areas without implantations. We have shown that not all atlases allow us to predict this change in the structure-function relationship within the brain.

Limitations. Our study is not without limitations. A major limitation is that we did not evaluate atlases in a diverse set of experimental systems, but rather limited our analysis to a contemporary topic in epilepsy linking two diverse measurement modalities of the brain to solve a clinical problem. We did not perform a feature selection analysis post-hoc to maximize ΔSFC at seizure onset; rather, we performed a comprehensive evaluation of many atlases to set a general framework and describe the nuances between the different atlases and their features. We hope this framework can be applied to many experimental designs. Ideally in our study, we required a whole-brain, volumetric atlas that covered the implanted SEEG electrode contacts. No such atlas existed. We opted for combining different atlases or developing randomly parcellated atlases used in previous publications [31,50], however, no general framework existed to determine which atlas should be used or clearly outlined the feature space of these atlases. We had no formal basis for how changing an atlas could change our results and eventual goal for translating network models to better treat epilepsy patients.

Another limitation is that we assume a change in SFC supports the hypothesis that seizures harness the underlying structural connectome of the brain (along with support from prior literature [23,24,51]). We may be biasing our results to select an atlas that maximizes ΔSFC. However, we wish to select a methodology that allows us to measure *any change* in brain state that accompanies seizure onset (explanatory validity), permitting us to probe epilepsy biology and understand the processes that govern seizure spread.

Finally, our analysis relies on the assumption that an atlas approach must be used to quantify SFC and does not consider an atlas agnostic approach nor if such an approach is appropriate. To study SFC using networks, both structure and functional networks must have nodes representing the same entity – neuroanatomical structures. The atlases defining anatomical structures (whether they are functionally, histologically, genetically, procedurally, multi-modally, or randomly defined) are the link between structural connectivity and functional connectivity measurements of the brain. To study SFC, we must rely on the neuroanatomical structures defined by an atlas, then localize electrodes to these regions and correlate the structural measurements (e.g. streamlines, fractional anisotropy, mean diffusivity) with functional measurements (e.g. cross-correlation, coherence, mutual information). Fundamentally, we are defining the nodes of the brain in advance,

which can alter our results; a more comprehensive discussion on defining the nodes of the brain are in Fornito et al. 2016 and Bijsterbosh et al. 2017[45,47].

In conclusion, the publication of atlases and their distribution across neuroimaging software platforms has risen exponentially over the last three decades. We simulate a study in which a researcher is blind to the development or features of an atlas and chooses one based on the availability in common neuroimaging pipelines and software (e.g. Freesurfer, DSI studio, FSL, SPM, QSIprep, fMRIprep, MRIcron, ANTs, and others). We advocate that while using a minimum of two atlases (one in the main text and one in the supplement) is one solution to understanding how results are affected by atlases choice, our framework provides a general solution. Researchers should instead justify *why* the atlas selected is appropriate using our framework above. Our study illustrates the critical need to evaluate the reproducibility of neuroscience research using atlases published alongside tools and analysis pipelines already established in the neuroscience community. Please see our GitHub for the atlases curated in this study along with their direct primary sources listed in Table S1. Our work provides a comprehensive resource for others investigating the brain's structure and function.

Funding

This work was supported by National Institutes of Health grants 1R01NS099348, K23-NS073801, 1R01NS085211, and 1R01MH112847. We also acknowledge support by the Thornton Foundation, the Mirowski Family Foundation, the ISI Foundation, the John D. and Catherine T. MacArthur Foundation, the Sloan Foundation, and the Paul Allen Foundation.

Competing Interests

The authors declare no competing interests.

Materials and Methods

Human Dataset. MRI data was collected from forty-one individuals, including thirteen healthy controls and twenty-eight drug-resistant epilepsy patients at the Hospital of the University of Pennsylvania. Twenty-four patients underwent stereoelectroencephalography (SEEG) implantation and four underwent electrocorticography (ECoG) implantation. Ten of the SEEG patients had clinically annotated seizures and were used for SFC analyses. Inclusion criteria consisted of all individuals who agreed to participate in our research scanning protocol, and (if they had implantations) allowed their de-identified intracranial EEG (iEEG) data to be publicly available for research purposes on the International Epilepsy Electrophysiology Portal (https://www.ieeg.org)[52,53]. Seizure evaluation was determined via comprehensive clinical assessment, which included multimodal imaging, scalp and intracranial video-EEG monitoring, and neuropsychological testing. This study was approved by the Institutional Review Board of the University of Pennsylvania, and all subjects provided written informed consent prior to participating. See Table S2 for subject demographics.

Structure. Methods and pipelines for structural connectivity generation and analysis are described in the following sections. Specific GitHub files and code are included where applicable.

Imaging Protocol. Prior to electrode implantation, MRI data were collected on a 3T Siemens Magnetom Trio scanner using a 32-channel phased-array head coil. High-resolution anatomical images were acquired using a magnetization prepared rapid gradient echo (MPRAGE) T1-weighted sequence (repetition time = 1810 ms, echo time = 3.51m, flip angle = 9, field of view = 240mm, resolution = 0.94x0.94x1.0 mm3). High Angular Resolution Diffusion Imaging (HARDI) was acquired with a single-shot EPI multi-shell diffusion-weighted imaging (DWI) sequence (116 diffusion sampling directions, b-values of 0, 300, 700, and 2000s/mm2, resolution = 2.5x2.5x2.5 mm3, field of view = 240mm). Following electrode implantation, spiral CT images (Siemens) were obtained clinically for the purposes of electrode localization. Both bone and tissue windows were obtained (120kV, 300mA, axial slice thickness = 1.0mm)

Diffusion Weighted Imaging (DWI) Preprocessing. HARDI images were subject to preprocessing pipeline QSIPrep to ensure reproducibility and implementation of the best practices for processing of diffusion images[54]. Briefly, QSIPrep performs advanced reconstruction and tractography methods in curated workflows using tools from leading software packages, including FSL, ANTs, and DSI Studio with input data specified in the Brain Imaging Data Structure (BIDS) layout.

Structural Network Generation. DSI-Studio (http://dsi-studio.labsolver.org, version: December 2020) was used to reconstruct the orientation density functions within each voxel using generalized q-sample imaging with a diffusion sampling length ratio of 1.25[55]. Deterministic whole-brain fiber tracking was performed using an angular threshold of 35 degrees, step size of 1mm, and quantitative anisotropy threshold based on Otsu's threshold[56]. Tracks with length shorter than 10mm or longer than 800mm were discarded, and a total of 1,000,000 tracts were generated per brain. Deterministic tractography was chosen based upon prior work indicating that deterministic tractography generates fewer false positive connections than probabilistic approaches, and that network-based estimations are substantially less accurate when false positives are introduced into the network compared with false negatives[31].To calculate structural connectivity, atlases listed in Table 1 were used. Structural networks were generated by computing the number of streamlines passing through each pair of structural regions in each specific atlas. Streamline counts were log-transformed and normalized to the maximum streamline count, as is common in prior studies[26,57–59]. GitHub: packages/imaging/tractography/tractography.py

Atlases. Atlas descriptions and sources used in this study are found in Table S1. All atlases were sourced in MNI space and if not already, resliced to dimensions 182x218x182. Atlases were linear then non-linear registered to T1w subject space using the ICBM 2009c Nonlinear Asymmetric template[60] and FSL flirt and fnirt. In addition to published standard atlases demarcating neuroanatomical and functional boundaries, we used whole-brain random atlases. A limitation of most standard atlases is that they may not have anatomical definitions for all regions of the brain, and therefore, implanted electrodes may not be assigned properly to a region. Whole-brain random atlases, in contrast, provide coverage to all implanted electrodes. They also allow for the ability to change some morphological properties (i.e. parcellation size), while keeping other morphologies the same (i.e. parcellation shape; Fig. 2d). A limitation of random atlases is that regions may not represent true anatomical or functional boundaries. With the limitations for each approach in mind, analyses were conducted for both standard and random atlases. Random atlases were built in the ICBM template space and covered all voxels, excluding those labeled as CSF or outside the template. To fill these points, a pseudo grassfire algorithm was applied[31]. Briefly, N points representing the number of ROIs of the atlas were randomly chosen as seed points. These seed points were iteratively expanded in all six Cartesian directions until all points were covered by one of the initial N seeds. After each iterative step, the smallest volume region expanded first. Random atlases created were of N equal to 10, 30, 50, 75, 100, 200, 300, 400, 500, 750, 1000, 2000, 5000, and 10000 ROIs. Five permutations for each N were created. GitHub code to generate random atlases: packages/imaging/randomAtlas/randomAtlasGeneration.py

Atlas Morphology: Volume and Sphericity. Atlas morphological measurements included regions of interest (ROI) size and shape, and

were measured with volume and sphericity calculations, respectively (Fig. 2). Region volume was calculated as the number of voxels in an ROI and log10 transformed. Region sphericity was calculated as the ratio of the surface area of a sphere with an equal volume of the ROI to the actual surface area of the atlas ROI. Under this definition, sphericity is bounded from 0 to 1 where 1 is a perfect sphere. For reference, a perfect cube and a hemi-sphere have a sphericity of 0.8 and 0.7 respectively. GitHub: packages/imaging/regionMorphology/regionMorphology.py

Structural Network Measures. We characterized the structural network topology of 52 atlases (Fig. 3 and Fig. S3). To quantify network topology, we examined density, mean degree, mean clustering coefficient, characteristic path length, and small worldness. Connectivity matrices were first binarized and a distance matrix was computed. The distance of any nodes that were disconnected from the main graph was set to the maximum distance between any pair of nodes in the main graph. Density, mean degree, clustering coefficient, and characteristic path length were then calculated on the binary, undirected graphs. Small worldness was calculated as the σ-ratio where $\sigma = \gamma/\lambda$ and is the ratio of the average, normalized clustering coefficient, C, to the normalized characteristic path length, I. γ = CG/CR and λ = lG/lR where G is the graph of interest and R represents a 'random' graph that is equivalent to G. To approximate the equivalent random graph R due to intractable computational costs[61], a well-known analytical equivalent CR = d/N and IR = log N/log d were used, where d denotes average nodal degree. All network measures were calculated using the Brain Connectivity Toolbox for Python. GitHub: papers/brainAtlas/Script_05_structure_02_network_measures.py

Function. Methods and pipelines for functional connectivity generation and analysis are described in the following sections. Specific GitHub files and code are included where applicable.

Intracranial EEG Acquisition. Stereotactic Depth Electrodes were implanted in patients based on clinical necessity. Continuous SEEG signals were obtained for the duration of each patient's stay in the epilepsy monitoring unit. Intracranial data was recorded at either 512 or 1024 Hz for each patient. Seizure onset times were defined by the unequivocal onset[62]. Seizure types were classified using ILAE 2017 criteria[63] as focal aware, focal impaired awareness, or focal to bilateral tonic-clonic. All annotations were verified and consistent with detailed clinical documentation. If a patient had more than one seizure annotated, the first seizure longer than 30 seconds without artifacts was used. iEEG times used in the study are sourced in the Box folder (below) and iEEG snippets are downloaded using GitHub script in packages/eeg/ieegOrg/downloadiEEGorg.py

Electrode Localization. In-house software[64] was used to assist in localizing electrodes after registration of pre-implant and post-implant neuroimaging data. All electrode coordinates and labels were saved and matched with the electrode names on IEEG.org. All electrode localizations were verified by a board-certified neuroradiologist (J.S.). Electrode coordinates in patient T1w space were assigned to an atlas ROI also registered in patient T1w space. Electrodes that fell outside the atlas of interest were excluded from subsequent analysis. GitHub: packages/atlasLocalization/atlasLocalization.py

Functional Connectivity Network Generation. Functional connectivity networks were generated from four periods: interictal, preictal, ictal, and postictal. (1) The interictal period consisted of the time approximately 6 hours before the ictal period. (2) The preictal period consisted of the time immediately before the ictal period. (3) The ictal period consisted of the time between the seizure unequivocal onset and seizure termination. (4) The postictal period consisted of the time immediately after the ictal period. Interictal, preictal, and postictal periods were 180 seconds in duration. Following removal of artifact-ridden electrodes, SEEG signals inside either GM or WM for each period were common-average referenced to reduce potential sources of correlated noise[65]. Next, each period was divided into 2s time windows with 1s overlap[66–69]. To generate a functional network representing broadband functional interactions between SEEG signals (Fig. 4b), we carried out a method described in detail previously[23,68]. Namely, signals were notch-filtered at 60 Hz to remove power line noise, low-pass and high-pass filtered at 127 Hz and 1Hz to account for noise and drift, and pre-whitened using a first-order autoregressive model to account for slow dynamics. Functional networks were then generated by applying a normalized cross correlation function ρ between the signals of each pair of electrodes within each time window, using the formula:

$$\rho_{xy} = \max_{\tau}\left[\frac{1}{T}\sum_{\tau=1}^{T}\frac{[x_k(t)-\bar{x}_k]*[y_k(t+\tau)-\bar{y}_k]}{\sigma_{x_k}\sigma_{x_y}}\right]$$

where x and y are signals from two electrodes, k is the 2s time window, t is one of the T samples during the time window, and τ is the time lag between signals, with a maximum lag of 0.5 s. Functional connectivity measurements were also calculated for coherence, zero time-lag Pearson and Spearman rank correlations with associated p-values, and mutual information. They are included in the freely available open source dataset but were not used in hypothesis testing in the study. Also freely available are the functional connectivity measurements in defined frequency bands reviewed in Newson and Thiagarajan 2019[70]. Networks are represented as full-weighted adjacency matrices. GitHub Code: GitHub: code/tools/echobase.py

Structure-Function Correlation. To quantify the relationship between structure and function in the epileptic brain, we computed the Spearman Rank correlation coefficient between the edges of the structural connectivity networks and the edges of the functional connectivity network (Fig. 4c). In the case where multiple electrodes fell in the same atlas ROI, a random electrode was selected to represent the functional activity of that neuroanatomically defined region. To reproduce these results, random seed was set to 42 using the NumPy Python package. Note that atlases with very small ROI volumes included more electrodes for SFC calculation. Electrodes that did not localize to an atlas were excluded from analysis. To average the SFC for all patients and each atlas (Fig. 5), SFC time-series was resampled to 100 seconds for each period and each sample was averaged together. GitHub code: packages/eeg/echobase/echobase.py

rsSFC and ΔSFC. Resting-state SFC (rsSFC) was defined as the SFC during the interictal period, approximately 6 hours before the ictal period. The mean SFC of that period was computed. ΔSFC was defined as the change in the mean SFC from the preictal to the ictal period (Fig. 5 top left panel). rsSFC and ΔSFC was calculated for each atlas (Fig. 6).

Statistics. Preictal and ictal SFC for each atlas were compared and significance was determined using the non-parametric repeated measures Wilcoxon signed-rank test. Bonferroni correction was applied over the 52 tests performed, equaling to the number of atlases studied.

Data availability and Reproducibility. All code files used in this manuscript are available at https://github.com/andyrevell/revellLab. All de-identified raw and processed data (except for patient MRI imaging) are available for download on Box. Link provided on GitHub. The GitHub repository used to analyze the data is also contained within Box. Raw imaging data is available upon reasonable request from Principal Investigator K.A.D.; tractography files generated from the imaging data are readily available on Box. iEEG snippets used specifically in this manuscript are contained within the Box data folder, while full iEEG recordings are publicly available at https://www.ieeg.org. The Python environment for the exact packages and versions used in this study in contained in the environment directory within the GitHub. The QSIPrep docker container was used for DWI preprocessing.

ACKNOWLEDGMENTS. We thank Adam Gibson, Carolyn Wilkinson, Jacqueline Boccanfuso, Magda Wernovsky, Ryan Archer, Kelly Oechsel, members of Andrew's Thesis Committee, and all other members and staff of the Center for Neuroengineering and Therapeutics for their continued help and support in this work.

References

[1] Klein, A. & Tourville, J. 101 labeled brain images and a consistent human cortical labeling protocol. *Frontiers in Neuroscience* **6** (2012).

[2] Mandal, P. K., Mahajan, R. & Dinov, I. D. Structural brain atlases: design, rationale, and applications in normal and pathological cohorts. *Journal of Alzheimer's disease : JAD* **31 Suppl 3**, S169–88 (2012).

[3] Dickie, D. *et al.* Whole brain magnetic resonance image atlases: A systematic review of existing atlases and caveats for use in population imaging. *Front Neuroinform* **11**, 1 (2017).

[4] Bohland, J. W., Bokil, H., Allen, C. B. & Mitra, P. P. The brain atlas concordance problem: Quantitative comparison of anatomical parcellations. *PLoS ONE* **4**, e7200 (2009).

[5] Beal, D. S. *et al.* The trajectory of gray matter development in broca's area is abnormal in people who stutter. *Frontiers in Human Neuroscience* **9** (2015).

[6] Van Horn, J. D. *et al.* Mapping connectivity damage in the case of phineas gage. *PLoS ONE* **7**, e37454 (2012).

[7] Barker, F. G. Phineas among the phrenologists: the american crowbar case and nineteenth-century theories of cerebral localization. *Journal of Neurosurgery* **82**, 672–682 (1995).

[8] Mazziotta, J. *et al.* A probabilistic atlas and reference system for the human brain: International consortium for brain mapping (icbm). *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **356**, 1293–1322 (2001).

[9] Evans, A. C., Janke, A. L., Collins, D. L. & Baillet, S. Brain templates and atlases. *NeuroImage* **62**, 911–922 (2012).

[10] Society, N. G. atlas. *National Geographic Society* (2011).

[11] Society, N. G. border. *National Geographic Society* (2011).

[12] Makris, N. *et al.* Decreased volume of left and total anterior insular lobule in schizophrenia. *Schizophrenia Research* **83**, 155–171 (2006).

[13] Hammers, A. *et al.* Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. *Human Brain Mapping* **19**, 224–247 (2003).

[14] Thomas Yeo, B. T. *et al.* The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology* **106**, 1125–1165 (2011).

[15] Schaefer, A. *et al.* Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity mri. *Cerebral Cortex* **28**, 3095–3114 (2018).

[16] Glasser, M. F. *et al.* A multi-modal parcellation of human cerebral cortex. *Nature* **536**, 171–178 (2016).

[17] Coalson, T. S., Van Essen, D. C. & Glasser, M. F. The impact of traditional neuroimaging methods on the spatial localization of cortical areas. *Proceedings of the National Academy of Sciences* **115**, E6356–E6365 (2018).

[18] Wu, Z., Xu, D., Potter, T., Zhang, Y. & The, A. D. N. I. Effects of brain parcellation on the characterization of topological deterioration in alzheimer's disease. *Frontiers in Aging Neuroscience* **11**, 113 (2019).

[19] Van Essen, D. C. *et al.* The wu-minn human connectome project: an overview. *NeuroImage* **80**, 62–79 (2013).

[20] Li, G. *et al.* Optimal referencing for stereo-electroencephalographic (seeg) recordings. *NeuroImage* **183**, 327–335 (2018).

[21] Freeman, W. J., Rogers, L. J., Holmes, M. D. & Silbergeld, D. L. Spatial spectral analysis of human electrocorticograms including the alpha and gamma bands. *Journal of Neuroscience Methods* **95**, 111–121 (2000).

[22] Slutzky, M. W. *et al.* Optimal spacing of surface electrode arrays for brain–machine interface applications. *Journal of Neural Engineering* **7**, 026004 (2010).

[23] Shah, P. *et al.* Characterizing the role of the structural connectome in seizure dynamics. *Brain : a journal of neurology* (2019).

[24] Ashourvan, A. *et al.* Pairwise maximum entropy model explains the role of white matter structure in shaping emergent co-activation states. *Commun Biol* **4** (2021).

[25] Proix, T., Bartolomei, F., Guye, M. & Jirsa, V. K. Individual brain structure and modelling predict seizure propagation. *Brain : a journal of neurology* **140**, 641–654 (2017).

[26] Wirsich, J. *et al.* Whole-brain analytic measures of network communication reveal increased structure-function correlation in right temporal lobe epilepsy. *NeuroImage. Clinical* **11**, 707–718 (2016).

[27] Willner, P. The validity of animal models of depression. *Psychopharmacology* **83**, 1–16 (1984).

[28] Bassett, D. S., Zurn, P. & Gold, J. I. On the nature and use of models in network neuroscience. *Nature Reviews Neuroscience* **19**, 566–578 (2018).

[29] Association, A. P. Technical recommendations for psychological tests and diagnostic techniques. *Psychological bulletin* **51**, 1–38 (1954).

[30] Salehi, M. *et al.* There is no single functional atlas even for a single individual: Functional parcel definitions change with task. *NeuroImage* **208**, 116366 (2020).

[31] Zalesky, A. *et al.* Whole-brain anatomical networks: does the choice of nodes matter? *Neuroimage* **50**, 970–983 (2010).

[32] Stacey, W. *et al.* Emerging roles of network analysis for epilepsy. *Epilepsy Res.* **159**, 106255 (2020).

[33] Cocchi, L. *et al.* Disruption of structure-function coupling in the schizophrenia connectome. *NeuroImage. Clinical* **4**, 779–787 (2014).

[34] Sathian, K. & Crosson, B. Structure-function correlations in stroke. *Neuron* **85**, 887–889 (2015).

[35] Gorgolewski, K. J. *et al.* Neurovault.org: a web-based repository for collecting and sharing unthresholded statistical maps of the human brain. *Frontiers in neuroinformatics* **9**, 8 (2015).

[36] Lawrence, R. M. *et al.* Standardizing human brain parcellations. *Scientific data* **8**, 78 (2021).

[37] Alexander, B. *et al.* A new neonatal cortical and subcortical brain atlas: the melbourne children's regional infant brain (m-crib) atlas. *NeuroImage* **147**, 841–851 (2017).

[38] Brennan, B. P. *et al.* Use of an individual-level approach to identify cortical connectivity biomarkers in obsessive-compulsive disorder. *Biological psychiatry. Cognitive neuroscience and neuroimaging* **4**, 27–38 (2019).

[39] Cabezas, M., Oliver, A., Lladó, X., Freixenet, J. & Cuadra, M. B. A review of atlas-based segmentation for magnetic resonance brain images. *Computer methods and programs in biomedicine* **104**, e158–77 (2011).

[40] Caspers, S., Eickhoff, S. B., Zilles, K. & Amunts, K. Microstructural grey matter parcellation and its relevance for connectome analyses. *NeuroImage* **80**, 18–26 (2013).

[41] Diedrichsen, J., Balsters, J. H., Flavell, J., Cussans, E. & Ramnani, N. A probabilistic mr atlas of the human cerebellum. *NeuroImage* **46**, 39–46 (2009).

[42] Belzung, C. & Lemoine, M. Criteria of validity for animal models of psychiatric disorders: focus on anxiety disorders and depression. *Biology of mood & anxiety disorders* **1**, 9 (2011).

[43] Association, A. E. R., Association, A. P., Education, N. C. o. M. i., (U.S., J. C. o. S. f. E. & Testing, P. *Standards for Educational and Psychological Testing* (American Educational Association, 2014).

[44] Sporns, O., Tononi, G. & Kötter, R. The human connectome: A structural description of the human brain. *PLoS computational biology* **1**, e42 (2005).

[45] Fornito, A., Zalesky, A. & Bullmore, E. *Fundamentals of Brain Network Analysis* (Academic Press, 2016).

[46] Sporns, O. The human connectome: a complex network. *Annals of the New York Academy of Sciences* **1224**, 109–125 (2011).

[47] Bijsterbosch, J., Smith, S. M. & Beckmann, C. F. *An Introduction to Resting State FMRI Functional Connectivity* (Oxford University Press, 2017).

[48] Yaakub, S. N. *et al.* On brain atlas choice and automatic segmentation methods: a comparison of maper & freesurfer using three atlas databases. *Scientific reports* **10**, 2837 (2020).

[49] Shmueli, G. To explain or to predict. *Statistical Science* **25** (2010).

[50] Mišić, B. *et al.* Cooperative and competitive spreading dynamics on the human connectome. *Neuron* **86**, 1518–1529 (2015).

[51] Betzel, R. F. *et al.* Structural, geometric and genetic factors predict interregional brain connectivity patterns probed by electrocorticography. *Nature Biomedical Engineering* (2019).

[52] *A multimodal platform for cloud-based collaborative research* (IEEE, 2013).

[53] Kini, L. G., Davis, K. A. & Wagenaar, J. B. Data integration: Combined imaging and electrophysiology data in the cloud. *NeuroImage* **124**, 1175–1181 (2016).

[54] Cieslak, M. *et al.* Qsiprep: An integrative platform for preprocessing and reconstructing diffusion mri. *Biorxiv* (2020).

[55] Fang-Cheng, Y., Wedeen, V. J. & Tseng, W.-Y. I. Generalized q-sampling imaging. *IEEE Transactions on Medical Imaging* **29**, 1626–1635 (2010).

[56] Otsu, N. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics* **9**, 62–66 (1979).

[57] Bonilha, L., Gleichgerrcht, E., Nesland, T., Rorden, C. & Fridriksson, J. Gray matter axonal connectivity maps. *Frontiers in Psychiatry* **6** (2015).

[58] Park, B., Eo, J. & Park, H.-J. Structural brain connectivity constrains within-a-day variability of direct functional connectivity. *Frontiers in Human Neuroscience* **11**, 408 (2017).

[59] Taylor, P. N. *et al.* The impact of epilepsy surgery on the structural connectome and its relation to outcome. *NeuroImage. Clinical* **18**, 202–214 (2018).

[60] Fonov, V. *et al.* Unbiased average age-appropriate atlases for pediatric studies. *NeuroImage* **54**, 313–327 (2011).

[61] Maslov, S. Specificity and stability in topology of protein networks. *Science* **296**, 910–913 (2002).

[62] Litt, B. *et al.* Epileptic seizures may begin hours in advance of clinical onset: a report of five patients. *Neuron* **30**, 51–64 (2001).

[63] Fisher, R. S. *et al.* Operational classification of seizure types by the international league against epilepsy: Position paper of the ilae commission for classification and terminology. *Epilepsia* **58**, 522–530 (2017).

[64] Azarion, A. A. *et al.* An open-source automated platform for three-dimensional visualization of subdural electrodes using ct-mri coregistration. *Epilepsia* **55**, 2028–2037 (2014).

[65] Ludwig, K. A. *et al.* Using a common average reference to improve cortical neuron recordings from microelectrode arrays. *Journal of neurophysiology* **101**, 1679–1689 (2009).

[66] Kramer, M. A. *et al.* Coalescence and fragmentation of cortical networks during focal seizures. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **30**, 10076–10085 (2010).

[67] Khambhati, A. N. *et al.* Dynamic network drivers of seizure generation, propagation and termination in human neocortical epilepsy. *PLoS computational biology* **11**, e1004608 (2015).

[68] Khambhati, A. N., Davis, K. A., Lucas, T. H., Litt, B. & Bassett, D. S. Virtual cortical resection reveals push-pull network control preceding seizure evolution. *Neuron* **91**, 1170–1182 (2016).

[69] Khambhati, A. N. *et al.* Recurring functional interactions predict network architecture of interictal and ictal states in neocortical epilepsy. *eNeuro* **4**, ENEURO.0091–16.2017 (2017).

[70] Newson, J. J. & Thiagarajan, T. C. Eeg frequency bands in psychiatric disorders: A review of resting state studies. *Frontiers in Human Neuroscience* **12**, 521 (2019).