Supplementary Online Content

Pierpaoli C, Nayak A, Hafiz R, et al; NIH AHI Intramural Research Program Team. Neuroimaging findings in US government personnel and their family members involved in anomalous health incidents. *JAMA*. doi:10.1001/jama.2024.2424

This supplementary material has been provided by the authors to give readers additional information about their work.

SUPPLEMENTARY APPENDIX

The NIH Neuroimaging Study of United States Government Personnel and their Family Members Involved in Anomalous Health Incidents

TABLE OF CONTENTS **PAGE**

Abbreviation Column1	Abbreviation Column2
White Matter (WM), Gray Matter (GM) and Cerebrospinal Fluid (CSF)	Primary Visual (PVIS) and Higher Visual (HVIS)
Estimated Intracranial Volume (eTIV) and Total Intracranial Volume (TIV)	Auditory (AUD), Language (LANG) and Precuneus (PREC)
Multi-cONtrast brain STRipping method (MONSTR)	Basal Ganglia (BG) and Sensorimotor Network (SMN)
Coefficient of Variation (CV)	Ventral Default Mode Network (VDMN)
Regional Analysis of Volumes Examined in Normalized Space (RAVENS)	Anterior Commissure-Posterior Commissure (AC-PC)
Multi-atlas region Segmentation utilizing Ensembles of registration algorithms and parameters (MUSE)	Mann Whitney (MW), Kruskal-Wallis (KW), Family Wise Error (FWE) and Benjamini-Hochberg (BH)
Tolerably Obsessive Registration and Tensor Optimization Indolent Software Ensemble (TORTOISE)	T1 weighted (T1W), T2 weighted (T2W), Echo time (TE), Repetition time (TR) and Inversion time (TI)
Diffusion Tensor Imaging (DTI)	Bandwidth (BW)
Fractional Anisotropy (FA)	Gradient Recalled Echo (GRE)
Mean Diffusivity (MD)	Gradient Moment Nulling (GMN)
Parenchymal Mean Diffusivity (Par-MD)	Intracranial Volume (ICV)
Radial Diffusivity (RD)	Echo Planar Imaging (EPI)
Axial Diffusivity (AD)	Quality Control (QC)
Cerebrospinal Fluid Signal Fraction (CSF-SF)	Baseline Visit (BV)
Non-Gaussianity (NG)	John Hopkins University (JHU)
Threshold Free Clustering Enhancement (TFCE)	Mean apparent propagator MRI (MAPMRI)
Propagator Anisotropy (PA)	Quantitative Medical Imaging (QMI)
Return to Origin Probability (RTOP)	General Linear Model (GLM)
Return to Axis Probability (RTAP)	Bayesian Multilevel (BML)
Return to Plane Probability (RTPP)	Montreal Neurological Institute (MNI)
Region of Interest (ROI)	Statistical Analysis Plan (SAP)
Analysis of Functional NeuroImages (AFNI)	National Institute of Mental Health (NIMH)
Deformable Registration via Attribute Matching and Mutual-saliency Weighting (DRAMMS)	National Institute of Biomedical Imaging and Bioengineering (NIBIB)
Percentile-based Score (Pscore)	Signal-To-Noise Separation (SNS)
Principal Component Analysis (PCA)	Signal-To-Noise Ratio (SNR)

Principal Components (PC)

Anomalous Health Incidents (AHI)

Magnetic Resonance Imaging (MRI)

National Institutes of Health (NIH)

NIH MRI Research Facility (NMRF)

Diffusion MRI (dMRI)

Functional Connectivity (FC)

Department of State (DOS)

Functional Magnetic Resonance Imaging (fMRI)

Susceptibility Weighted Imaging (SWI)

Functional Balance Disorder (FBD)

Persistent Postural-Perceptual Dizziness (PPPD)

Blood Oxygenation Level Dependent (BOLD)

Resting State fMRI (RS-fMRI)

Resting State Network (RSN)

Anterior Salience (ASAL)

Posterior Salience (PSAL)

Visuo-spatial (VSPL)

Post-Traumatic Stress Disorder (PTSD)

Left Executive Control Network (LECN)

Right Executive Control Network (RECN)

Posterior default mode network (PDMN)

Dorsal Default Mode Network (DDMN)

Susceptibility Weighted Imaging (SWI)

Beck Anxiety Inventory (BAI)

Healthy Volunteer (HV)

Inter-Quartile Range (IQR)

Blood Biomarker (Blood-BM)

Wechler Adult Intelligence Scale (WAIS)

Glial Fibrillary Acidic Protein (GFAP)

Neurofilament light chain (Nfl)

Cervical Vestibular Evoked Myogenic Potential (cVEMP)

Ocular Vestibular Evoked Myogenic Potential (oVEMP)

Subjective Visual Vertical (SVV)

Sensory Organization Test (SOT)

Neurobehavioral Symptoms Inventory (NSI)

PTSD-Check List version 5 (PCL-5)

Fatigue Severity Scale (FSS)

Epworth Sleepiness Scale (ESS)

Beck’s Depression Inventory II (BDI-II)

Patient Health Questionnaire-15 (PHQ-15)

Satisfaction with Life Scale (SWLS)

Pain Catastrophizing Scale (PCS)

Superior Longitudinal Fasciculus (SLF)

Corpus Callosum (CC)

Cerebellar Peduncle (CP)

External Capsule (EC)

Superior Fronto-Occipital Fasciculus (SFOF)

In addition to classic radiological evaluation, several advanced research MRI modalities are available for the assessment of brain structure and function. In this work, we employed volumetric structural MRI, diffusion MRI (dMRI), and resting state functional MRI (RS-fMRI). The following sections will expand on techniques and methods implemented for each of these modalities.

eAppendix 1. Structural Volumetric and Clinical MRI

Volumetric structural MRI relies on high resolution acquisitions of T1-weighted (T1W) and T2-weighted (T2W) MRIs. From these images, brain tissue is segmented into gray matter (GM), white matter (WM), and cerebral spinal fluid (CSF) compartments, assuming specific signal properties for these different tissue types. Following segmentation, the volume of these compartments can be calculated both regionally and for the whole brain. Volumetric MRI is an established technique for detecting brain atrophy following normal aging and various pathologies. However, volumetric MRI is susceptible to motion artifacts and results can be unreliable in presence of pathologies that significantly alter the normal signal tissue properties of different brain tissue types (e.g., tumors). In the following sub-sections, we highlight the various methods and techniques involving T1W anatomical MRI used in this study.

1.1 Image Acquisition

On the Siemens Prisma, equipped with a 32-channel head coil, the structural MRI sequences included the acquisition of fat-suppressed T2W TSE and T1W MPRAGE images. The scanning parameters for the T1W image were: sagittal acquisition, 1 mm^3 isotropic resolution, matrix size=176x256x256, TE/TR:3.3/2530ms, and for the T2W image: axial 2D turbo-spin-echo acquisition, 1.25x1.25x1.7 mm^3 resolution, matrix size=192x192x110, TE/TR: 72/8810ms, strong fat-suppression. In addition, for inter-scan reproducibility analysis of structural scans (see **eAppendix1.4.2**) T1W MPRAGE scans acquired on the clinical scanner were included. Imaging details for those scans are as follows: SIEMENS Biograph with 16 channel head and neck coil, sagittal acquisition, 1 mm^3 isotropic resolution, TE/TR=3.03/2530ms, FA=7 degs, TI=1100ms.

Susceptibility Weighted Imaging (SWI) was also acquired on the SIEMENS Biograph clinical scanner. In each patient, T2*-weighted gradient recalled echo (GRE) magnitude and phase images were obtained using the vendor-provided SWI acquisition, which uses a standard 3D GRE pulse sequence with first-order gradient moment nulling (GMN) flow compensation. Contrast parameters for the GRE SWI were TR = 40 msec, TE = 25 msec, flip angle = 15°, and

bandwidth (BW) = 50 Hz/pixel. Geometric parameters were image matrix = 448 × 439 × 72, and voxel size = 0.5 × 0.5 × 2 mm. In addition, parallel imaging (GRAPPA) was used with an acceleration factor of 2. Phase enhanced images were automatically computed from the magnitude and phase images as part of the vendor software.

1.2 Preprocessing

ROI-wise volumetric measurements were obtained using the FreeSurfer[1] software with default parameters, however, the estimation of the total intracranial volume (ICV) was done using MONSTR[2]. MONSTR is more accurate and reproducible than FreeSurfer in extracting ICV, due to FreeSurfer's instability in the extraction of brain from surrounding dura. MONSTR utilizes information from T1/T2W data to more accurately discriminate brain tissue from surrounding non brain tissue giving a more accurate measurement of ICV. In addition, another set of data was generated by the RAVENS[3-5]-MUSE[6] pipeline. Finally, for ROI-wise analysis of the diffusion data we used the MUSE ROIs[6] for the subcortical regions and the cerebellum and a set of ROIs informed by the widely used John Hopkins University (JHU) white matter (WM) ROIs[7] but redrawn on a diffusion tensor average brain template obtained from data from the Human Connectome Project[8].

1.3 Brain Volume Measurement

FreeSurfer is an automated brain segmentation software that can provide volumetric information about cortical and subcortical structures of the brain. For our study, we used FreeSurfer version 6.0, a stable version of the software, with 'recon-all'. However, incorporation of an enhanced bias correction N4 algorithm is an useful feature from the new versions (7-7.4) that could add value to the current analysis. Bias field inhomogeneity correction is used to correct for potential signal inhomogeneities in T1W data. While FreeSurfer version 6.0 incorporates N3 bias field correction within its pipeline, N4 correction was implemented as it is more powerful[9]. It is essential that signal inhomogeneities across the T1W data are corrected using the most optimal algorithm since not fully accounting for this correction can result in misclassification of tissues affecting the accuracy of volume measurements. Especially, since we are evaluating scans from two different scanners, each scanner may have its own level of signal inhomogeneities that is introduced in the data that could further add to the variability measured between scanners.

The Freesurfer pipeline also performs brain masking and extraction within its pipeline. However, if the pipeline only uses T1W data, the brain extraction and masking tend to be

unstable, resulting in over or underestimation of the brain mask, due to the limited information in T1W data to distinguish between dura and GM. In addition, eTIV (estimated intracranial volume, the normalization variable output by FreeSurfer) can also be affected by inaccurate brain masking that occurs due to limited information from T1W data[10].

Therefore, we decided to incorporate a masking tool, Multi-cONtrast brain STRipping method (MONSTR)[2] that uses information from T1W and T2W data, within the FreeSurfer pipeline to aid in the brain masking procedure. We used the total intracranial volume (TIV) derived from the brain mask computed with MONSTR to normalize the regional volumes measured via the FreeSurfer pipeline. In summary, we provided the FreeSurfer algorithm, with the N4 bias corrected T1W data together with its respective MONSTR mask. FreeSurfer 'recon-all' command was used with the -3T tag to generate 49 segmented brain regions and volumes in the aseg.stats file. However, to reduce granularity we combined some of the smaller regions into larger regions as shown in **eTable 1**. Normalization was performed by dividing all measured volumes by the TIV computed from the MONSTR brain mask. In addition, we have eliminated from our analysis small or insignificant regions that are prone to variability due to their size.

1.4 Interscan and Inter-scanner Reproducibility

1.4.1 Interscan Reproducibility

Interscan reproducibility analysis was performed to evaluate the stability of research scans and to quantify any experimental error. Four healthy volunteers from the control cohort of the study were scanned 5 times, each, with minimum duration between visits ranging from 1-9 days (**eTable 2**). For the structural volumetric reproducibility analysis, we computed regional volume measurements using FREESRUFER-MONSTR pipeline as detailed in **eAppendix1.3**.

1.4.2 Inter-scanner Reproducibility

In this study, 110 subjects (85.3% of the study cohort, Controls = 36, AHI Participants = 74) have T1W and T2W scans acquired on both clinical (Siemens Biograph mMR) and research scanner (Siemens Prisma). Averaging the measured regional volumes from data acquired on two different scanners can provide us with an estimate of best representative measure of brain volume per subject, potentially eliminating the experimental confounds added due to scanner differences or segmentation errors. However, before considering the data combination between the two scanners, we assessed the inter-scanner data reproducibility to make sure the data was highly reproducible across scanners. For the 19 subjects with unavailable data (16 from Siemens Biograph mMR + 3 from Siemens Prisma) due to absence of T1W data, we performed

a linear regression with available data from one scanner to predict the missing values in the other scanner, after we established that the available volumetric data were highly reproducible between scanners. Since the Siemens Prisma had fewer missing values, in the first step, the Siemens Biograph mMR was considered as the response variable (Y), while the Siemens Prisma scanner was considered as the predictor variable (X). The coefficients from the linear regression using the available data between the two scanners was used to predict the missing values in the data from the Siemens Biograph mMR. In the second step, this order of operation was reversed and the missing values from the Siemens Prisma were predicted (Y) using the Siemens Biograph mMR data as the predictor (X). Consequently, the volumetric data for each of the 129 participants across two scanners were averaged and normalized by their respective averaged TIV, to be used in the ROI wise statistics.

1.5 Volumetric Analysis Using RAVENS-MUSE

Verma et al[11] used Regional Analysis of Volumes Examined in Normalized Space (RAVENS)[3-5] and Multi-atlas region Segmentation utilizing Ensembles of registration algorithms and parameters (MUSE)[6] brain segmentation and volume measurement pipeline where they reported global differences between participants and controls. To perform a comparative volumetric analysis with our data we used the RAVENS-MUSE pipeline and analysis steps, detailed by Verma and colleagues. Briefly the RAVENS-MUSE pipeline is described as follows: MUSE anatomical labeling approach was used to segment T1W scans of each subject into GM, WM, and CSF regions. The RAVENS-MUSE maps were computed in template space by warping each subject's GM, WM, and CSF using the Deformable Registration via Attribute Matching and Mutual-saliency Weighting (DRAMMS)[12] transformation generated from registering each subject to the John Hopkins University (JHU) template.[13,14] The RAVENS output is a map of intensity values in template space that are directly proportional to the corresponding regional volume measured in native space. To control inter-subject variability, the RAVENS-MUSE maps were normalized by the TIV of each subject. The statistical analysis performed, and results obtained using the RAVENS-MUSE maps are detailed in **eAppendix3**.

eAppendix2. Structural Diffusion MRI

The dMRI technique is exquisitely sensitive to microstructural and organizational features of the brain. From diffusion MRI acquisitions it is possible to compute quantitative metrics that characterize different aspects of the probabilistic diffusion displacement profile of water molecules present in the tissue. To detect changes in brain tissue composition, microstructure, and white matter organization, dMRI is probably the most powerful neuroimaging technique. However, dMRI measurements are highly susceptible to acquisition artifacts that can severely impair the biological significance of the findings; therefore, ensuring accuracy and reproducibility of the results is fundamental when using this technique. The following sub-sections expand on the various methods involved with dMRI in the current study.

2.1 Image Acquisition

The dMRI data were collected using an in-plane parallel imaging factor of 2, no simultaneous multi-slice, isotropic voxel resolution of 2 mm^3 and field-of-view of 220x220x180 mm^3. The entire dMRI dataset consisted of data acquired with four phase-encoding directions: Anterior-Posterior (AP), Posterior-Anterior (PA), Right-Left (RL), Left-Right (LR) to enable correction of geometric distortions. It has been shown in the literature[15] that this 4-way phase encoding direction scheme can increase reproducibility with respect to other phase encoding direction schemes. Acquisitions with the same phase-encoding orientation (AP/PA and RL/LR) employed identical diffusion experimental designs, each with near-optimal gradient distributions sampling the entire diffusion sphere to minimize the effects of imaging gradients and eddy-currents. The two differing sets of diffusion gradients (AP/PA vs RL/LR) were designed to be complementary. Diffusion sensitization for each phase-encoding direction acquisition covered a range of b-values (0 s/mm^2 (n=2), 50 (n=4), 300 (n=2), 1100 (n=16), 2500 (n=16)) for a total of 160 volumes. The entire diffusion experimental design is reported in **eTable 5**. For all acquisitions, identical timing parameters were used (TE/TR: 71/9800ms), which were obtained from the optimal parameters of the RL/LR phase-encoded b=1100s/mm^2 acquisition with no partial-Fourier. To achieve identical timings, b=2500s/mm^2 acquisitions employed 7/8 partial k-space sampling. A square field of view was used for all acquisitions to have identical echo train length for both AP/PA and RL/LR phase encoding. The scan time for the dMRI part was 31 minutes and 9.5 minutes for the anatomical images for a total of 40.5 minutes.

2.2 Preprocessing

The dMRI data were processed using the TORTOISE pipeline.[16,17] As per the pipeline, each subject's manually AC-PC'ed[18] T2W 1 mm^3 resampled data was used as a reference in the processing and DTI data was processed with an 1 mm^3 isotropic output voxel resolution. The four different phase encoding datasets were all used to obtain optimal corrections of image distortions[15] and were combined for further computation of dMRI metrics. Diffusion Tensor Imaging (DTI) metrics, such as Mean Diffusivity (MD=1/3 of the Trace of the diffusion tensor), Fractional Anisotropy (FA), Axial (AD) and Radial (RD) diffusivities, were computed from data up to the bvalue of 1100 s/mm^2. Mean apparent propagator metrics (MAP MRI)[19] including Non-Gaussianity (NG), Propagator Anisotropy (PA), Return to the origin probability (RTOP), Return to the axis probability (RTAP), and Return to the plane probability (RTPP) were computed from the entire dMRI dataset including the bvalue of 2500 s/mm^2. Finally, the Cerebrospinal-Fluid Signal Fraction (CSF-SF) and the Parenchymal Mean Diffusivity (Par-MD) were computed from a dual compartment model[20] from data up to the bvalue of 1100 s/mm^2. The availability of intermediate b-values between 0 to 1100 s/mm^2 ensured that the estimation of these parameters was not ill-posed.

A study template was generated from the individual diffusion tensors of all subjects included in the study using the methodology proposed by Irfanoglu et al[21]. All voxel-wise and ROI-wise analyses were performed in the space defined by the study template. The study template creation and ROI definition and its value extraction is explained in **eAppendices 2.3** and **2.4** below. A T1W study template was also created by voxel-wise averaging of the individual T1W images after applying the same spatial warping computed from registering DTI data, per subject, to the study template.

2.3 Study Template Creation

Voxel-wise group analysis requires a voxel-wise correspondence between data included in the analysis. This is achieved by either registering the data to a publicly available template or by creation of a study specific template. In addition, a scalar or tensor-based registration can be used in the generation of a template. We use a tensor-based registration[21] to create the study template at 1 mm^3 isotropic voxel resolution, as tensor based registration has been shown to be more accurate than scalar based registration[21] in registering white and gray matter structures.

A study template was generated in December 2021, using diffusion tensors from all the subjects included in the study and the template was not updated for removal/addition of 10 subjects. This study template was only used as a target space to perform the group analysis. For the subjects included in the analysis, their diffusion tensors (48 controls and 81

participants) were re-registered to the study template tensor using a no-smoothing setting in DRTAMAS.[21] The affine and diffeomorphic transformation generated by registering each subject diffusion tensor (DT) to the study template was used to bring the respective DTI and MAPMRI metrics, into the study template space.

2.4 Region of Interest (ROI) Definition

2.4.1 White Matter ROIs from Connectome Template

In neuroimaging research, the set of JHU white matter (WM) ROIs[22] is widely used. While these ROIs are defined with high regional accuracy, they are prone to some right and left structural instabilities.[23] In addition, to transfer the JHU ROIs from the JHU template to study template, a scalar registration needs to be performed that can further introduce some instabilities due to mis-registrations[21]. Therefore, we used white matter ROIs, which are informed by the JHU ROIs and are defined on a connectome DT template.[24] For our analysis we registered the connectome DT template to the study template DT using a tensor-based registration.[21] Thereafter, the ROIs from the connectome template were transferred onto the study template using ANTS scalar transformation tool with nearest neighbor interpolation.[25] White matter ROI information were extracted in study template space from each subject's DTI and MAPMRI metrics using ITK-SNAP.[26] **eFigure 1** shows the anatomical location of the WM ROIs in the study template.

2.4.2 Additional ROIs Defined from MUSE Brain Segmentation

For cerebellum WM and GM vermal lobules, and GM nuclei regions, we have adopted the MUSE[6] ROIs. One study found that MUSE ROIs are more reliable in the subcortical structures than other segmentation methods.[27]

The following procedure was used to incorporate cerebellum WM, GM, and GM nuclei, in our analysis. For the RAVENS-MUSE pipeline, as detailed in eAppendix1.3, MUSE segmentation was performed on T1W data. To transfer the MUSE ROIs from T1W onto the respective DTI data, T1W images were rigidly registered to the T2W AC-PC reoriented data. The spatial transformations generated from the rigid registration with ANTS[25] of T1W data to AC-PC T2W data, was applied to the MUSE ROIs in T1W space and using nearest neighbor interpolation in ANTS[25] were warped into the AC-PC T2W space. Given DTI data was oriented in its respective final AC-PC T2W space (see **eAppendix2.2**), the MUSE ROIs defined on T1W data were now transferred onto the respective DTI space and values for select cerebellum WM, GM and GM nuclei were extracted from each subject's DTI metrics. **eFigure 2** shows the location of select GM nuclei, and cerebellum ROIs.

2.5 Interscan Reproducibility

DTI data was acquired on the four healthy volunteers (see **eAppendix1.4.1**) and each baseline visit was processed as detailed in **eAppendix2.2**. However, for repeated visits, each diffusion data was processed in TORTOISE with its respective T2W data that was rigidly registered using ANTS[25] to the baseline visit's AC-PC T2W data. DT was computed per visit and a subject specific DT template was created using DRTAMAS.[21] Each subject template DT was then registered using DRTAMAS to the study specific template.

WM ROIs were defined on each subject's repeated DTI data as follows:

1) Spatial transformations generated for each subject's repeated DT during the subject template creation was combined with the spatial transformation generated from registering the subject's template to the study template.
2) The combined spatial transformation was applied to each visit's DTI scalar metric bringing them into the study template space that has the WM ROIs.

GM nuclei, cerebellum WM, and cerebellum lobule ROIs from MUSE ROIs were defined on each subject's repeated DTI data as follows:

1) MUSE ROIs defined on baseline T1W data of the subject were transferred into the baseline DTI data as detailed in **eAppendix2.4.2**.
2) Baseline DT from subject was registered to the study template DT as detailed in **eAppendix2.3**.
3) Spatial transformations generated from registration of baseline DT to study template DT were applied to scalar transform the MUSE ROIs from baseline DTI into the study template using nearest neighbor interpolation.[25]
4) Spatial transformations generated for each subject's repeated visit DT during the subject template creation was combined with the spatial transformation generated from registering the subject's template to the study template.
5) The inverse of the combined transformation generated in step 4, per visit, was applied using nearest neighbor interpolation[25] to the MUSE ROIs in study template. The MUSE ROIs were transferred onto each visit's DTI data.

Quantitative values were derived from each WM and GM ROIs, per scalar metric and scan.

eAppendix3. Statistical Analysis Plan (SAP)

Given the variability in clinical presentation, timing and modalities of the acute events experienced, the uncertainties on mechanism, spatial extent, and regional distribution of the injury, we concluded that the study involving the anomalous health incidents[28] (AHIs) had to be essentially exploratory. Instead of focusing on specific regions, the analyses had to cover the entire brain and we needed to explore the results of several metrics from different MRI modalities. An exploratory study obviously requires adjustment for multiple comparisons to control for false positive findings. However, when the number of exploratory questions is large (multiple variables and regions), multiple comparison adjustment can be highly penalizing by introducing false negatives. Given that both false positive and false negative findings would be detrimental for understanding the mechanism of AHIs and clinical management of the participants affected, we incorporated a comprehensive analysis plan involving multiple statistical method. This includes conventional analysis both with and without multiple comparison adjustments and Bayesian Multilevel (BML)[29] modeling which intrinsically addresses the issue of multiplicity in the conventional model.

We performed both voxel-wise and region of interest (ROI)-wise analyses for the structural (T1-weighted and diffusion) MRI data. For resting state functional MRI (RS-fMRI) data, we did a network-based analysis with ROIs that span the majority of the gray matter. For group level assessments, we applied a conventional non-parametric modeling approach for the voxel-wise analyses and both conventional non-parametric and Bayesian modeling approaches for the ROI-wise analyses. We adopted a BML framework (see **eAppendix3.3.1**) to incorporate the entire set of ROIs in a single model with the above-mentioned goal of overcoming the multiplicity issue and to provide a more robust estimation of the effects at the regional level.

The statistical analyses plan (SAP) was developed together with the team of statisticians at the Quantitative Medical Imaging (QMI) lab at the National Institute of Biomedical Imaging and Bioengineering (NIBIB) and the Scientific and Statistical Computing Core team at the National Institute of Mental Health (NIMH) of the National Institutes of Health (NIH). The SAP was modified in February 2022, with the evolution and characterization of the AHI severity (AHI1 and AHI2).[30] As a result, in addition to the original analysis with the two-group comparison that included the entire group of participants with AHI vs. controls, we also performed a three-group comparison, which involved AHI subgroups (AHI1, AHI2) and controls.

At the whole brain level, general linear models (GLMs) and related contrasts were first prepared in FSL[31] for the two-group and three-group voxel-wise comparisons. The model parameters were then passed to the non-parametric 'randomise' module in FSL to perform 10000 permutations and compute family wise error (FWE) corrected P value maps with threshold free clustering enhancement (TFCE)[32] for an unadjusted threshold of $P < 0.01$ (two-sided). At the ROI level, the conventional statistical analysis plan for the two-group comparison involved the Mann-Whitney *U* test, which is a non-parametric alternative to a two-sample *t*-test, at an unadjusted threshold of $P < 0.05$ (two-sided). For the three-group comparisons, we performed a non-parametric Kruskal-Wallis analysis of variance with *post-hoc* pair-wise comparisons using Mann-Whitney *U* tests, at an unadjusted threshold of $P < 0.05$. The rationale for selecting non-parametric approaches was the skewed nature of the distribution of several brain imaging derived quantities used in this study. For the ROI-wise analyses, the adjustment for multiple comparisons was performed using the Benjamini-Hochberg (BH)[33] method.

Although we recruited additional controls to ensure the sample groups were more age-matched, for all voxel-wise analyses age was added as a covariate of no interest. For the ROI-wise analyses, we observed a significant effect of age across the population in several imaging metrics and ROIs. To account for the age effect, we used a quantile regression approach to fit the median ($\tau = 0.5$) age of the entire sample population which is described in more detail in **eAppendix3.3.2.**

An important part of our analysis plan was to come up with a strategy to investigate individual differences/patterns of brain imaging metrics, while also highlighting any systematic differences across the groups in each ROI. Although Z-scores are popularly used to show how a subject deviate from a normative distribution, it can introduce large imbalance in extreme values when the data is heavily skewed[34,35]. We employed a percentile-based approach[34] to compute a score, called 'Pscore', which accounts for the skewness in the data and help attain a closer fit to a normal Z-distribution (see **eAppendices 3.3.3** and **3.3.4**). 'Pscore' incorporates an individual's position in the left/right tail of the normative distribution and normalizes that individual's deviation from the normative median by the corresponding length between the normative median and the 5th/95th percentile edge values of the normative distribution, respectively. We used these 'Pscores' to generate the heatmaps shown in our manuscript and the Supplementary Appendix. The heatmaps simultaneously inform the individual deviations from the control distribution and the systematic differences across ROIs in a single figure.

The diffusion study produced eight quantitative metrics, some from the diffusion tensor (DTI) and some from the mean apparent propagator (MAPMRI) analysis. Some of these metrics are expected to be intrinsically correlated. Therefore, we performed a principal component analysis (PCA) (**eAppendix3.3.5**) using the Pscores from each of these diffusion metrics. The strategy was to reduce the dimensions of several correlated variables and to obtain two meaningful orthogonal components that explain the most variance in the data, obtaining a comprehensive but concise picture of the diffusion properties in white matter.

We strove to reinforce any statistical results with a measurable effect of interest, which makes the results more interpretable and meaningful[36]. To better interpret the scale by which changes occurred in participants with AHI compared to controls, we also computed the difference in median (%), which computes by what percentage the difference in median between the participants with AHI and controls changed with respect to the median in the control group. Moreover, where applicable, we report median, interquartile ranges and/or standardized effect estimates and their confidence intervals. We also report both unadjusted and BH-adjusted P values for multiplicity. For the Bayesian models, we report posterior probabilities higher than 0.95 for positive effects and lower than 0.05 for negative effects for each pairwise contrast.

The final part of our SAP involved evaluating the relationship (see **eAppendix6**) between brain imaging derived quantities and the various clinical parameters depicted in our companion paper. We performed Spearman's rank correlation for the correlation between the clinical and brain imaging variables. To further assess the relationship between a few specific clinical measures and brain-imaging derived metrics, a simple linear regression model was used. All ROI-wise statistical analyses were performed in R statistical software.[37]

The following sections expand on the statistical analyses performed on the structural volumetric and dMRI data and describes the results associated with each analysis. Since the T1W and dMRI are both structural imaging modalities, the metrics derived using these techniques underwent the same statistical tests in turn. We first show the strong 'interscan' and 'inter-scanner' reproducibility of these two imaging modalities and then present results from whole brain voxel-wise and ROI-wise analyses, respectively.

3.1 Evaluation of Reproducibility for Volumetric and Diffusion Measurements

3.1.1 Interscan Reproducibility for Volumetric MRI

We calculated a coefficient of variation (CV) of structural volumetric data per ROI per subject, by dividing the *standard deviation* of volume across 5 visits with the *average* volume across 5

visits. This ratio was then multiplied by 100 to represent CV in terms of percentages (%). **eTable 3** shows the average CV across the 4 subjects, showing high reproducibility in volumetrics for large ROIs: Total Intracranial Volume (ICV)=0.3%, Total Parenchyma= 0.5%, Total GM=0.6% and Cerebral WM=0.6%.

3.1.2 Inter-scanner Reproducibility for Volumetric MRI

Reproducibility analysis was performed between the clinical and research scanner data, by using linear regression of the volume measurements from regional data detailed in **eAppendix1.3**. A simple linear regression model was used to explain the variance in the clinical scanner by using the research scanner (PRISMA) as the explanatory variable. **eTable 4** reports the slope and intercept of the linear regression model along with their respective confidence intervals. The correlation between the inter-scanner variables and their R^2 values for the volumetric results in different ROIs are also provided. More than 51% of the ROIs demonstrated an R^2 value > 0.90 and ~86% of the ROIs had an R^2 value > 0.70. About 14% of the ROIs showed R^2 values < 0.70, primarily, the right Pallidum, bilateral Accumbens area and the bilateral Cerebellar WM ROIs. Therefore, while the inter-scanner reproducibility is excellent for global ROIs (Total Parenchyma, Total GM and Cerebral WM, etc.) some smaller structures that often suffer from signal loss in MRI and the highly folded, complex cerebellar structures demonstrated comparatively lower reproducibility across scanners.

3.1.3 Interscan Reproducibility Analysis of Diffusion Data

For the dMRI data of the four subjects, CV averaged across all four subjects was calculated per ROI as described in **eAppendix3.1.1**. **eTable 6** shows excellent overall reproducibility for dMRI metrics. The median (Q1, Q3) CV % across all WM ROIs show very high reproducibility: MD=0.6 (0.5, 0.7)%, FA= 0.6 (0.5, 0.8)%, AD=0.5 (0.4, 0.6)%, RD=1.0 (0.8, 1.2)%. The MAPMRI metrics also show very high reproducibility: RTOP=1.4 (1.2, 1.6)%, RTAP=1.4 (1.1, 1.6)%, RTPP=0.3 (0.3, 0.4)%, PA=0.2 (0.1, 0.4)% and NG=1.5 (1.0, 2.0)%. The worst reproducibility is observed for the CSF-SF: CV=7.4 (5.6, 9.5)%; and Par-MD: CV=1.7 (1.2, 2.8)%. These measures are noisy across all the ROIs as the number of volumes in the intermediate b-value shell is low, making it not very reliable in the extraction and measurement of CSF-SF and hence Par-MD.

3.2 Two Group Comparison Voxel-wise

In the two-group comparison, controls (n=48) were compared against all participants with AHI (n=81). This involves the conventional approach, where each voxel is considered independent and undergoes independent statistical testing. Typically, some form of multiple comparison

adjustment is warranted. Following we detail our voxel-wise two-group comparisons, the adjustment method used and effect magnitude computation.

3.2.1 Whole Brain Voxel-wise Group Analysis

A non-parametric FSL[31] randomise threshold free clustering enhancement (TFCE)[32] analysis was used for voxel-wise group analysis. A general linear model (GLM) was set up for two group comparison using demeaned age as a covariate and with 10000 permutations. The analysis was performed to output maps that are both unadjusted and TFCE adjusted for whole brain voxel-wise multiple comparisons. In addition, a whole brain voxel-wise analysis with no TFCE was performed on data generated by the RAVENS[3-5]-MUSE[6] pipeline.

The group wise whole brain voxel wise analysis was run on the following metrics in the study space: FA, MD, Par-MD, CSF-SF, RD, AD, NG, PA, RTAP, RTOP, RTPP, and RAVENS MUSE GM and WM volumetric maps. The results for whole brain voxel wise analysis are shown in **eFigure 3**.

3.2.2 Magnitude Map Computation

Magnitude of difference between the participants with AHI and controls, per DTI, MAPMRI and volume maps were created as follows: $(median_{\text{participants}} - median_{\text{controls}})/median_{\text{controls}} \times 100$. It is essentially equivalent to the difference in median (%) values shown for the ROIwise analysis. Results for percent magnitude change are shown in **eFigure 3**.

3.3 Two Group Comparison ROI-wise

3.3.1 ROI Analysis Using Bayesian Multilevel (BML) Modeling

We adopted a Bayesian multilevel (BML) modeling approach[29] to explore the region-level differences between the two groups. Specifically, with y_{ij} as the effect of the *i*th region (each of DTI or T1w metrics) from the *j*th subject in the *k*th group, we formulated the following model:

$$y_{ij} \sim N(\mu_{ij}, \sigma^2);$$

$$\mu_{ij(k)} = \beta_0 + \beta_1 \text{sex}_j + \beta_2 \text{age}_j + \beta_3 \text{group}_j + \theta_{0i} + \theta_{1i} \text{sex}_j + \theta_{2i} \text{age}_j + \theta_{3i} \text{group}_j + \tau_j;$$

$$(\theta_{0i}, \theta_{1i}, \theta_{2i}, \theta_{3i})^T \sim N(\mathbf{0}_{4\times1}, \boldsymbol{S}_{4\times4}), \tau_j \sim N(0, \pi^2);$$

$$i = 1, 2, \ldots, k, j = 1, 2, \ldots, n, k = 1, 2;$$

where β_0, β_1, β_2, and β_3 are the effects associated with the overall intercept, sexes, age, and groups; θ_{0i}, θ_{1i}, θ_{2i}, and θ_{3i} represent the *i*th region's effects that correspond to the overall intercept, sexes, age, and groups; τ_j codes the effect from the *j*th subject; σ^2 is the distributional

variance for the likelihood; $\boldsymbol{S}_{4\times4}$ is the variance-covariance matrix for the region-level effects θ_{0i}, θ_{1i}, θ_{2i}, and θ_{3i}; and π^2 is the distributional variance for subject-level effects τ_j.

The inferences for group differences at each region were assessed through the overall posterior distribution. In contrast to the conventional modeling methodology, in which each region is handled separately, the above model incorporates the data hierarchy with information efficiently shared and regularized across regions in a single model. Therefore, instead of treating multiple testing adjustment as a post hoc step under the traditional statistical framework, we adjust the region-level effects through partial pooling to improve model quality as well as predictive accuracy.[29,38] The region-level posterior distribution was obtained from Markov Chain Monte Carlo simulations using the *R* package *brms* with the default settings of weakly informative hyperpriors.[39] The adoption of such priors is intended to learn from the data without the injection of strong prior information for the following three types of parameters: (a) population effects or location parameters (e.g., intercept and slopes), (b) standard deviations or scaling parameters for lower-level (e.g., regions, subjects) effects, and (c) various parameters such as the covariances in a variance-covariance matrix. Noninformative hyperpriors are adopted for population-level effects. In contrast, weakly-informative priors are utilized for standard deviations of lower-level parameters such as varying intercepts and slopes at the subject and region level, and such hyperpriors include a Student's half-t(3,0,1) or a half-Gaussian $N_+(0, 1)$ (a normal distribution with restriction to the positive side of the respective distribution). For variance-covariance matrices, the LKJ correlation prior[40] is used with the shape parameter taking the value of 1 (i.e., jointly uniform over all correlation matrices of the respective dimension). Lastly, the standard deviation σ for the residuals utilizes a half Cauchy prior with a scale parameter depending on the standard deviation of the input data. The consistency and full convergence of the Markov chains were confirmed through the split statistic $\hat{R}$ being less than 1.1.[38] The effective sample size (or the number of independent draws) from the posterior distributions based on Markov Chain Monte Carlo simulations was more than 200 so that the quantile (or compatibility) intervals of the posterior distributions could be estimated with reasonable accuracy.

eTable 7 summarizes the results of the ROI analysis using the BML modeling approach for the two-group comparison. For conciseness, we report in the table only results for metrics which have a posterior probability higher than 0.95 for positive effects and lower than 0.05 for negative effects. The Bayesian approach incorporates the entire set of ROIs in a single model, quantifying a probability of how likely one group is different from the other. For

example, the Genu of Corpus Callosum (CC) has a $P^{+} = 0.003,$ for RTOP, which means the likelihood of the controls having higher RTOP than participants in this ROI is $< 0.3\%$.

3.3.2 Quantile Regression Fit of Median Age

To account for age effects in the region of interest (ROI)-wise analysis, we incorporated a quantile regression[41,42] approach. The rationale behind using this approach is rooted in several factors. First, noise is assumed to be random and normally distributed and secondly variance is assumed to be constant in, typical mean regression and parametric Z-score calculations. This implies that the data is normally distributed and the variance in the quantitative variable remains constant across age, sex and other related demographics in the population. If one defines statistical significance based on Pvalues and confidence intervals, a violation of the normality and constant variance assumption can lead to inaccurate inferences. Sherwood B. et al.,[43] had addressed some of these issues and used a quantile regression modeling of percentiles for neuropsychological test scores. We have recently shown that imaging data with high skewness can introduce bias[34,35] if the analysis involves Z-score computation with respect to a normative distribution. We found that the normative distribution of multiple dMRI derived metrics can be heavy tailed, leading to bias and imbalance[34,35] in extreme positive and negative Z-scores. In a recent study,[34] we proposed a novel percentile based approach to correct for this imbalance, which is discussed later in more detail (see **eAppendices 3.3.3** and **3.3.4)**. The median value is typically a more robust central tendency estimate in such heavy tailed distributions. Therefore, we developed an in-house script entirely based on R[37] that performs age correction using quantile regression to fit the median overall sample age of the two groups. The *'quantreg'* package was used and the correction was contingent on the age effect being significant at $P < 0.05$. We describe the steps and procedures of this method below in more detail.

Before performing the quantile regression, the covariate 'age' was centered by the median age across the two groups. Mathematically, this involves subtracting the median age from every participant's age. This provides a more meaningful interpretation of the 'intercept' which would now represent the response variable e.g., the DTI metric at the median age of 41 years instead of 0 years. A 'tau' of 0.5 was chosen to obtain the median age fit. The residuals from the regression model were further shifted by the intercept value at the median age of 41 years. Importantly, not all linear fits would be significant. Therefore, a conditional algorithm was written to only perform age adjustment, if the regression model was significant at $P < 0.05$, to avoid unnecessary ubiquitous regression across all ROIs, which may introduce spurious adjustments and overfitting.

In **eFigure 5**, we visualize the quantile regression results for a representative ROI (Superior Cerebellar Peduncle Left) for two different DTI metrics (FA in **eFigure 5A-B** and PA in **eFigure 5C-D**). The aim was to demonstrate the significant linear trend in age across two groups and their distributions before and after removing the age effect. The conventional mean regression model fit is also shown (solid blue line) to compare with the quantile regression model (solid red line). On the left column in **eFigure 5A** and **C** shows the regression models for FA and PA, respectively, along with the slope of age and the P value for significance of the model fit. **eFigure 5B** and **D** represent the FA and PA trends and values after removing the age effect. The scatter plot values in **eFigure 5B** and **D** represent the residuals shifted by the intercept value, which, now represents the DTI metrics at the median age of the two samples, 41 years. On the right column, the histogram distributions show that within the same ROI, age effects tend to be more skewed for PA and slightly heavy tailed for FA. Therefore, adopting a quantile regression model is more reliable and accurate under these conditions.

3.3.3 Computing Percentiles

To assess the scale by which an individual can deviate from the control distribution, we applied a percentile-based approach[34] to generate metrics that can be interpreted as somewhat analogous to classical Z-scores. As a necessary first step, we computed subject-wise percentile rank values within each ROI using the following formula –

$$\boldsymbol{p_{rank_{ij}} = \frac{n_{C \le x_{ij}}}{n_C} \times 100; \quad i = 1, 2, 3 \ldots N, \;\; j = 1, 2, 3, \ldots M} \qquad [1]$$

In equation [1], $\boldsymbol{x_{ij}}$ represents a DTI or structural volume metric of the i^{th} individual for the j^{th} ROI, where $\boldsymbol{i}$ sequentially represents $\boldsymbol{1 \ldots N}$ participants from the control group and those with with AHI ($\boldsymbol{N = n_C + n_{AHI}}$) and $\boldsymbol{j}$ represents $\boldsymbol{1 \ldots M}$ ROIs, respectively. It is important to note that the value, $\boldsymbol{n_{C \le x_{ij}}}$represents the number of subjects within the "control" group having a value equal to or less than $\boldsymbol{x_{ij}}$. Also, to note, the denominator $\boldsymbol{n_C}$ represents the total number of subjects only in the "control" group. Therefore, every individual's percentile rank is computed with respect to the sample size of the control group. Naturally, it is possible that an AHI participant may have an $\boldsymbol{x_{ij}}$ that is less than the minimum value in the control group. In such cases, the values are set to the 0th percentile.

Computing the percentile values with respect to the control group is significant because, we would later compute by how many units can an individual deviate with respect to the difference between the median and the 5th and 95th percentile edges of the "control" distribution. With $\boldsymbol{n_C = 48}$, the smallest resolution within the control group can be 2.08 ≈ the

2nd percentile. As a result, we computed the ~5th percentile edge being closest to the ~6th percentile (6.25 to be exact). Likewise, for the ~95th percentile edge, the closest was the ~95th percentile (95.82 to be exact).

3.3.4 Computing Pscores

As mentioned earlier, we opted to generate a percentile-based metric, called a *Pscore*,[34,35] that indicates how far a subject deviate from the control distribution with respect to the difference between the median and 5th/95th percentile edges in the control group. The interpretation is somewhat similar (but not identical) to a standardized Z-score which indicates how many standard deviations a subject is away from the mean. The *Pscore* indicates by how many units a subject is away from the median of the control group with respect to the difference between the median and either of two extreme left/right percentile edges in the control distribution. We call this metric a *Pscore* as this is computed based on percentile ranks described in the earlier section. The interpretation of directions in Z-scores is that a +3 is indicative of a subject 3 standard deviations above the mean and contrarily, a -3 would be 3 standard deviations below the mean. Either way, the subject is typically considered extreme if the values are above +3 or below -3. Similarly, for Pscores, a '+3' would indicate that a subject is above the median in controls by 3 units of the difference between the median and the 95th percentile edge in the control distribution. On the other hand, a '-3' would indicate that a subject is below the median in controls by 3 units of the difference between the median and the 5th percentile edge in the control distribution. We formalize these concepts below to depict how the Pscores are computed for the two tails:

$$\boldsymbol{P_{5_{ij}}} = (-)\mathbf{1.645}\ \frac{\boldsymbol{d_{x_{ij}}}}{\boldsymbol{d_5}};\quad \begin{cases} \boldsymbol{d_{x_{ij}}} = \boldsymbol{x_{ij}} - \boldsymbol{m_{C_j}} & \boldsymbol{x_{ij}} < \boldsymbol{m_{C_j}} \\ \boldsymbol{d_5} = \boldsymbol{m_{C_j}} - \boldsymbol{x_{Cj_5}} & \boldsymbol{x_{ij}} < \boldsymbol{m_{C_j}} \end{cases} \qquad [2]$$

$$\boldsymbol{P_{95_{ij}}} = (+)\mathbf{1.645}\ \frac{\boldsymbol{d_{x_{ij}}}}{\boldsymbol{d_{95}}};\quad \begin{cases} \boldsymbol{d_{x_{ij}}} = \boldsymbol{x_{ij}} - \boldsymbol{m_{C_j}} & \boldsymbol{x_{ij}} > \boldsymbol{m_{C_j}} \\ \boldsymbol{d_{95}} = \boldsymbol{x_{Cj_{95}}} - \boldsymbol{m_{C_j}} & \boldsymbol{x_{ij}} > \boldsymbol{m_{C_j}} \end{cases} \qquad [3]$$

Like equation [1], in equations [2] and [3], $\boldsymbol{i, j}$ and $\boldsymbol{x_{ij}}$ represent the same parameters. The numerator $\boldsymbol{d_{x_{ij}}}$ is the difference between $\boldsymbol{x_{ij}}$ and the median of the control group, $\boldsymbol{m_{C_j}}$ for the j^{th} ROI. If $\boldsymbol{d_{x_{ij}} < 0}$, it means the current subject has a value less than $\boldsymbol{m_{C_j}}$, thereby, falling on the left-hand tail between $\boldsymbol{m_{C_j}}$ and the 5th percentile edge value, $\boldsymbol{x_{Cj_5}}$ in the control group, and the denominator $\boldsymbol{d_5}$ is computed as the difference between $\boldsymbol{m_{C_j}}$ and $\boldsymbol{x_{C_{j5}}}$. On the other hand, if $\boldsymbol{d_{x_{ij}} > 0}$, it means the current subject has a value above the $\boldsymbol{m_{C_j}}$, therefore, falling on

the right-hand tail between $\boldsymbol{m_{C_j}}$ and the 95th percentile edge in the control group, $\boldsymbol{x_{Cj_{95}}}$ and the denominator $\boldsymbol{d_{95}}$ is computed as the difference between $\boldsymbol{x_{Cj_{95}}}$ and $\boldsymbol{m_{C_j}}$. Both $\boldsymbol{P_{5_{ij}}}$ and $\boldsymbol{P_{95_{ij}}}$ takes on the polarity of $\boldsymbol{d_{x_{ij}}}$, generating the negative and positive scores, respectively.

Please note, the scalar '1.645' represents the absolute Z-score value corresponding to the 5th and 95th percentiles of a normal distribution. This step brings the Pscores to the Z-score scale for better interpretations and make them comparable. For example, a Pscore of ± 2 would correspond to a Z-score of ± 2 from a normal distribution. The heatmaps shown in the manuscript and the supplementary appendix comprise these Pscores. Owing to the nature of the equation and like Z-scores, Pscores can have very large values, i.e., identifying highly extreme participants. As a result, we also capped the Pscores at ± 3 for the heatmaps. Therefore, any Pscore higher than $> +3$ and < -3 was assigned the darkest shades of red and blue in the heatmaps, respectively.

3.3.5 PCA Analysis

To comprehensively summarize the overall information from the dMRI metrics we incorporated a principal component analysis (PCA) in R[37] by adopting the 'FactoMineR'[44] package. All dMRI metrics except for Par-MD and CSF-SF were used for this analysis, because even with intermediate b-value acquisitions, these metrics were noisier than others. For each diffusion metric, the Pscores from all 50 WM ROIs were averaged for each individual and a global matrix containing *n* x *p* elements were generated, where *n = 1, 2, 3 … 129* subjects and *p = 1, 2, 3 … 8* diffusion metrics or 'variables' in the PCA terminology. By taking advantage of dimensional reduction, we reduced the metrics to two orthogonal components that retained ~ 87% of the variance in the data. **eFigure 6A** shows the level of inherent collinearity that exists among some of the diffusion metrics (upper panel) along with the corresponding quantitative Pearson's correlation, '*r'* values (lower panel). **eFigure 6B** shows the scree plot depicting the percentage of variance explained by the principal components (PCs). Since ~ 87% of the variability is explained by the first two PCs, we emphasized on these two components to summarize the data, pivoting on the global average across ROIs for each diffusion metric in a single biplot.

3.4 Controls vs. AHI Subgroups Comparison

3.4.1 Whole Brain Voxel-wise Analysis of AHI Subgroups

An additional analysis was performed dividing the participants with AHI into two subgroups based on the factors described in the executive summary from the Intelligence Community.[30]

All the statistical analyses detailed in **eAppendix3.2** were performed to evaluate the whole brain voxel-wise differences between the following group pairs: controls and AHI1; controls and AHI2. The randomise program was run after a General Linear Model (GLM) was set up in FSL to compare the three groups. Results of the analysis are shown in **eFigures 3** and **4**. In addition, ROI-wise analysis was run as detailed in **eAppendix3.3**.

3.4.2 ROI-wise Kruskal-Wallis Test of AHI Subgroups

In the main manuscript, we had demonstrated ROI-wise results for the Mann Whitney U test between controls and participants. To assess differences in dMRI quantitative measurements across controls, AHI1 and AHI2 sub-types, a Kruskal-Wallis (KW) omnibus test was run along with adjustments for multiple comparisons using the Benjamini-Hochberg[33] (BH) method. In **eTable 8**, we present differences in diffusion metrics across these groups in specific ROIs at an unadjusted level ($P < 0.05$) for the KW test and showing an effect of at least > 2% in any of the pair-wise group comparison tests. We also provide the corresponding BH-adjusted values for these KW tests along with BH adjusted Wilcoxon Rank Sum test results across all the group pairs.

We do not observe any specific group demonstrating consistent abnormalities across these ROIs to make any conclusive remarks on systematic differences. We can observe that within the right External Capsule (EC) and the right Caudate, both AHI1 and AHI2 groups demonstrate reduced CSF-SF and NG when compared to the control group. However, in the bilateral Superior Longitudinal Fasciculus (SLF) the AHI2 group demonstrates reduced FA and enhanced RD compared to both the control and AHI1 groups. On the other hand, in the bilateral Superior Fronto-Occipital Fasciculus (SFOF) the AHI1 group shows enhanced FA and reduced RD compared to the control and AHI2 groups. Similarly, the AHI1 group also demonstrates enhanced AD compared to the control and AHI2 groups in the bilateral Inferior Cerebellar Peduncle (CP), which also survived the BH adjustment at the omnibus level of the KW test (**eTable 8**). At an unadjusted level ($P < 0.05$), there were more ROIs that showed some difference albeit with very small effects, of order ~0-1%. Here, we have shown ROIs where the effect was at least > 2%. One can observe that the difference in percentage effect is quite small to indicate any substantial differences across these groups.

3.4.3 ROI-wise Pscore Heatmaps with AHI Subgroups

The heatmaps in the main paper show participants divided into controls and those with AHI. **eFigure 7A-I** show the heatmaps of all other MRI metrics not covered in the main manuscript. **eFigure 8A-L** present heatmaps from all MRI metrics by further subgrouping the participants

with AHI into AHI1 and AHI2 categories. Moreover, within each group, the participants are sorted by increasing interval after AHI.

3.4.4 Overall-WM PCA Analysis with AHI Subgroups

Similar to the biplot with the two groups (**eFigure 9.A**), the PCA biplot with the three groups is shown in **eFigure 9.B**. We can observe that the 95% confidence ellipse for the three groups largely overlap, indicating that the overall WM profiles in these groups are quite similar.

3.5 Longitudinal Assessment

Longitudinal changes in the volumetric and dMRI measurements were assessed with respect to time from the onset of AHIs. The median (Q1, Q3) time of delay from the day of AHI onset till the first MRI scan was 80 (36, 544) days with a range of 14-1505 days. We incorporated whatever data was available until a fixed date (November 7th, 2022). This included 49 participants with AHIs with longitudinal scans approximately 6-12 months apart having at least 2 visits, 17 participants with at least 3 visits, and 8 participants having 4 visits. The median (Q1, Q3) inter-scan interval was 371 (298, 420) days. Please refer to **eAppendix2.5** for the ROI definitions for the longitudinal assessment. We examined the ROIs from **Table1** in the main text, where we showed group level differences from the cross-sectional analysis at an unadjusted $P < 0.05$. In **eFigure 10A-G** we plotted the corresponding volumetric and diffusion MRI metrics with respect to the time after AHI (in days). The 'PAT' labels help locate an AHI participant and each progressive dot represents a follow up scan from the baseline visit (BV). Across time, most participants remain stable, with variability consistent with what is expected from the experimental errors as measured in the reproducibility portion of our study. At the individual level there does not appear to be evidence of a systematic trend of changes across visits, whether for subjects whose measurements at the first visit were very close to that of the median of the controls or for subjects whose measurements differed notably from the median of the controls. This indicates the absence of evolving lesions. At the population level there is no evidence of a dependency on the results from the time after AHIs.

eAppendix4. Resting State Functional MRI (RS-fMRI)

Blood Oxygenation Level Dependent (BOLD)[45] imaging measures local changes in brain hemodynamics which are reflective of brain activity. Resting state fMRI (RS-fMRI) is a BOLD based technique that measures how fMRI signal fluctuations are correlated in different brain areas in absence of external stimuli or tasks. The spatial pattern of synchronicity of resting state fMRI signal fluctuations has been found to be remarkably reproducible across individuals and populations and has been interpreted to be reflective of intrinsic functional networks present in the brain.[46-52] For this reason, synchronicity of resting state fMRI signal fluctuations has been traditionally referred to as "functional connectivity".[53,54] In the following sections we describe the various methods involved in the RS-fMRI imaging.

4.1 Acquisition

Gradient echo planar imaging (EPI) was used to collect 206 whole-brain functional volumes from 125 participants (controls = 45, participants with AHI = 80). RS-fMRI Data from 3 controls and 1 AHI participant were not available. In addition, during quality control (QC) assessment (see **eAppendix4.2** below), 11 participants (controls = 3, participants with AHI = 8) were removed due to excessive motion, leaving an effect sample size of 114 participants (controls = 42, participants with AHI = 72) The imaging parameters for the acquisition include TR = 2000 ms; TE = 27 ms; flip angle = 90°, 36 slices with slice thickness of 3.6 mm, matrix = 64x64; FOV = 220 x 220 mm^2 and acquisition voxel size = 3.438 x 3.438 x 3.6 mm^3. The subjects were instructed to remain still in the scanner and to avoid movement as much as possible with eyes fixed to a cross on an overhead screen.

4.2 Pre-Processing

The RS-fMRI data was pre-processed in AFNI[55,56] (using "afni_proc.py"). Briefly, the first four time points were removed to account for the delay in magnetic stabilization. The rest of the time series data was then 'despiked' to remove abnormally large spikes in the signal and slice-time corrected to bring all images to the same origin in time. Each functional image was motion corrected by first identifying the volume with the minimum outlier fraction and then registering all other volumes to it. FreeSurfer segmentation was run on the anatomical T1-weighted (T1w) images to obtain tissue specific masks using 'recon-all', specifically, for the lateral ventricles and white matter (for denoising). The T1-weighted images were then skull stripped and warped to the standard Montreal Neurological Institute (MNI) template using "@SSwarper". The skull stripped image was aligned to the corresponding subject's EPI base

volume using affine transformation. The transformations (across EPI, EPI to anatomical and anatomical to template) were all concatenated into a single matrix, to nonlinearly warp the EPI data to final MNI template space to minimize blurring due to re-gridding. These normalized images were then resampled to an isotropic voxel resolution of 3 x 3 x 3 mm^3. Note, for the ROI-wise analysis, no spatial smoothing was applied to the data to avoid blurring across non-contiguous regions and mixing in information from voxels outside of a given atlas ROI.

For the regression block, the first three principal components obtained from the lateral ventricle signals, the demeaned motion parameters and their derivatives were all added as covariates in a multiple linear regression (Multivariable) model to remove nuisance signals per voxel. Note, no bandpass filtering was applied, as it has been shown to over-penalize the degrees of freedom and introduce unrelated spurious correlations[57] and significantly reduce signal-to-noise separation (SNS) and test-retest reliability[58]. Finally, fast ANATICOR[59] was performed, creating voxel-wise regressors from local white matter (as segmented by FreeSurfer) to help correct for physiological and/or hardware artefacts. To account for excessive motion which may introduce spurious correlations,[60] we applied a motion censoring criterion, which censors volumes with large motion (Enorm> 0.3 mm) or with at least 5% outlier voxels in a whole brain mask.

In addition, single subject quality control was performed to evaluate a combination of the input data and successful completion of processing steps. This was performed using afni_proc.py's QC HTML and gen_ss_review_table.py.[61] As part of this procedure, participants were removed from further analysis if their data met any of the following criteria –

a) lost > 60% of the degrees of freedom (i.e., retains only < 40% of available data after censoring)
b) censor fraction > 0.3
c) averaged censored motion of 0.2
d) maximum displacement > 8 mm
e) global correlation (GCOR) > 0.15

4.3 ROI-wise Functional Connectivity (FC)

We implemented the 90 functional ROI atlas from Shirer et al.[62] for the ROI-wise analysis, to be consistent with the ROIs used in the study by Verma et al.[11] The ROIs were resampled to the same orientation and voxel resolution (3 x 3 x 3 mm^3) of the pre-processed data. Out of the 90, 4 ROIs were removed. Firstly, the 'Left Thalamus' from the primary visual network (PVIS)

comprised only 1 voxel, compared to the other constituent, 'Calcarine Sulcus', comprising 327 voxels. This large disparity in voxel sizes among constituents of the same network can introduce unrelated bias and variability and therefore, the smaller ROI was removed. Secondly, the other 3 ROIs – 'Right Lobule IX' (from ventral DMN (VDMN)), and Bilateral Lobule VIII (from Visuospatial (VSPL)), comprise lower inferior cerebellar lobules which suffer from low signal-to-noise ratio (SNR) and often do not get full coverage due to field of view limitations. Therefore, 86 ROIs were used to extract average time series from the preprocessed data. Pearson's correlation was performed between pairs of columns (ROIs) with the extracted time series per subject, generating an [86_{ROIs} x 86_{ROIs}] correlation matrix, which is symmetric and contains 86×85/2=3,655 unique values. The correlation values were then Fisher's Z transformed and the process repeated for all subjects, creating a [$114_{Subjects}$ x 86_{ROIs} x 86_{ROIs}] three-dimensional matrix for the full set of N = 114 subjects. Therefore, the [$114_{Subjects}$ x 86_{ROIs} x 86_{ROIs}] matrix was reshaped to a [$114_{Subjects}$ x 3655_{roi_pair}] two-dimensional matrix. These values were then corrected for age and a Mann-Whitney *U* test was performed for each ROI and adjusted for multiple comparisons using the BH method.

4.4 Within-Network Functional Connectivity (FC)

For the within-network analysis, only the subset of unique ROI-to-ROI connections belonging to a given network were used. For example, the Posterior Salience (PSAL) network consists of 12 ROIs, which therefore contains 12×11/2=66 unique connections within its correlation matrix. The average of these 66 Fisher's *Z* values were computed and converted back to an '*r*' value representing the within network functional connectivity of this network. This was repeated for all subjects and for all networks generating a [$114_{subjects}$ x $13_{networks}$] matrix. Note, the primary visual network (PVIS) had only one ROI and therefore no unique ROI-to-ROI correlations are possible for this network, hence, 13 out of the 14 networks were used for further analysis and visualizations.

eAppendix5 Statistical Analysis for RS-fMRI

For the three-group comparisons among Controls, AHI1 and AHI2, we first applied an omnibus non-parametric Kruskal-Wallis test to evaluate across group differences with multiple comparison adjustment for having 13 networks. Based on our central tendency plots (see **eFigure 11**), we observed that in most of the 13 networks, AHI1 group demonstrated a lower median value of within-network FC than both the Control and AHI2 groups. The Controls and AHI2 groups also had very similar within-network FC median values. Therefore, to reduce unnecessary testing, we followed up with *post-hoc* Mann-Whitney *U* tests for only the Control and AHI1 group comparisons adjusted for multiple testing using the BH adjustment approach. For both the two group and three group comparisons, we also applied a Bayesian multilevel modeling approach, which does not incur a multiple comparison penalty. We compared the results from the Bayesian model with those from the conventional model to find complementary evidence that the Control and AHI1 groups had significantly different within-network connectivity.

5.1 Group Level Differences in Functional Connectivity

5.1.1 Two Group Bayesian Analysis

eFigure 11 shows the box plots of within-network FC between the controls and participants with AHI. The purple lines with asterisks show which of these networks demonstrated evidence to the participants with AHI having enhanced FC compared to controls. PSAL showed the strongest evidence ($P^+ < 0.002$) where the participants with AHI demonstrate a < 0.2% posterior probability to have enhanced FC compared to controls. More evidence can also be found but with weaker probabilities ($P^+ < 0.021$ and 0.042) for the SMN and ASAL networks, where the participants with AHI show < 2.1% and < 4.2% posterior probability, respectively, to have enhanced FC compared to controls. Therefore, the strongest evidence from both our conventional and Bayesian analyses indicates that the average FC within the PSAL network may differ in the participants with AHI compared to controls.

5.1.2 Controls vs. AHI Subgroups Conventional Analysis

eFigure 12 shows the box plots with the within-network FC values across controls, AHI1 and AHI2 subgroups. The median of the distributions of participants in the AHI1 category tended to be slightly lower than either of the other two groups in majority of the RSNs, which was examined for statistical significance. The omnibus Kruskal-Wallis (KW) test revealed an overall difference in FC across the three groups within the ASAL and PSAL networks (P_{KW} = 0.005 and 0.011, respectively) for an unadjusted threshold of $P_{KW} < 0.025$. Evaluating

differences specifically between participants within AHI1 category and controls in *post-hoc* analysis with BH adjustments (at $\alpha = 0.05$) for multiple comparisons revealed, participants in AHI1 category demonstrate reduced FC compared to controls within the ASAL network ($P_{BH} = 0.021$) but the PSAL network fails to survive the $P_{BH} < 0.025$ by a small margin ($P_{BH} = 0.029$).

5.1.3 Controls vs. AHI Subgroups Bayesian Analysis

eFigure 12 also shows the results from the Bayesian analysis. The purple lines with asterisks indicate which of these networks show some evidence towards participants in the AHI1 category having enhanced average FC compared to controls. ASAL and PSAL showed the strongest evidence ($P^+ < 0.002$) where participants in the AHI1 category demonstrate a < 0.2% posterior probability to have enhanced FC compared to controls. More evidence can also be found but with weaker probabilities ($P^+ < 0.026$ and 0.028) for the SMN and BG networks, where participants within AHI1 category show < 2.6% and < 2.8% posterior probabilities, respectively, to having enhanced FC compared to controls. Therefore, the strongest evidence from both our conventional and Bayesian analyses indicates that the average FC within the ASAL and PSAL networks may differ in participants within the AHI1 category.

eAppendix6 Assessing Relationship of Imaging Metrics with Clinical Measures

In our companion paper, we discuss various clinical measures that were used to further evaluate the participants with AHI. A set of 41 clinical parameters were obtained from several domains, including the vestibular, balance, auditory, and cognitive behavior. These measures have been described in more detail in our companion paper by Chan and colleagues. Our goal for this analysis was to assess how much of the variance in the imaging derived quantities is explained by the clinical parameters within brain regions where participants with AHI showed some evidence of abnormalities. We also assessed the strength of these relationships in each group using quantile regression of the median quantity.

6.1 Correlation Between Imaging Metrics from Significant ROIs and Clinical Measures

In **Table 2** of the main manuscript, we presented some ROIs where one or more dMRI derived metrics had shown group level differences between controls and participants with AHI. To assess the relationship between diffusion metrics and clinical parameters (**eAppendix6** above) within these WM ROIs, we used the scores from the first two principal components (PCs) obtained from the PCA analysis (see **eAppendix3.3.5**) and MD for the GM ROIs. In addition, the ROIs from the structural volumetric MRI and the within network FC of networks from fMRI were also assessed. We used Spearman's method to obtain correlation values (ρ) for each of the 41 clinical variables to bypass parametric assumptions of normality. **eFigure 13** shows the complete correlation plot for the controls (top, above the middle horizontal line) and the participants with AHI (bottom, below the middle horizontal line). A few scattered correlations were significant at an unadjusted level ($P < 0.01$). The corresponding ρ values are shown in bold within the correlation plot in both sections for controls and participants, respectively. In general, the strength of correlation between imaging metrics and clinical parameters was weak and we did not observe any conclusive evidence to a maladaptive pattern.

6.2 Relationship of FC in Salience Networks with Anxiety and PTSD

Since we observed some differences in FC within the Salience networks (PSAL and ASAL), we further evaluated if the FC within these networks correlated with clinical measures of anxiety and PTSD. We used the Beck Anxiety Inventory[63] (BAI) for anxiety, particularly, the total scores that range from 0 to 63.[64] The PTSD-Check List Version 5[65,66] (PCL-5) scores were used for PTSD. We performed a linear regression analysis with the FC of each Salience network as the dependent variable (y) and the corresponding clinical score as the independent

or explanatory variable (x). **eFigure 15A-B** show the regression lines and corresponding equations for the participants with AHI and the subgroups (AHI1 and AHI2), respectively. The R^2 values, as well as the confidence intervals for the slope and intercept are provided in the figures. We did not observe any significant relationship between the FC within salience networks and the clinical parameters for anxiety and PTSD.

eAppendix7. Outcome Metrics

In this section, we describe the outcome metrics used to test the AHI and Control cohorts.

eAppendix7.1 Volumetric Outcome Metrics

From structural MRI data, segmentation of various brain regions and the extraction of the total intracranial volume (TIV) were performed for each subject as described in eAppendix1.3. The segmented regions are listed in eTable4. All volumes were extracted as absolute values in cubic millimeters (mm^3), but the volumetric outcome metric used for analysis is the Relative Volume computed as Volume divided by the respective TIV. This Relative Volume ranges from 0 to 1 and is a dimensionless quantity.

eAppendix7.2 Diffusion MRI Outcome Metrics

7.2.1 Diffusion MRI Introduction

Diffusion MRI allows the measurement of features of the molecular displacement of water molecules due to random thermal motion. In a jar of water, the probabilistic molecular displacement (also called diffusion propagator) is spherical because the probability of displacement is the same in all spatial directions, and the root mean square displacement is linearly related to the diffusion time through a scalar diffusion coefficient D. For free water at 37C, D is $3 * 10^{-3}$ mm^2/s In biological tissues, the diffusivity of water is generally lower than that of pure water, and the diffusion displacement profile may not be spherical. For example, in white matter water diffusivity is higher in the direction parallel to the fibers than perpendicular to them.

The importance of diffusion MRI for the assessment of biological tissues is that diffusion distances probed by water molecules in the typical time of the measurement are of the order of microns, making diffusion MRI exquisitely sensitive to microstructural features of the tissue such as cell density, cell morphology, spatial arrangement of cellular structures, volume, composition, and geometry of extracellular space, as well as membrane integrity and permeability.

7.2.2 Outcome Metrics Computed from the Diffusion Tensor

Diffusion tensor MRI (DTI) is the simplest techniques to extract basic quantitative information from the measured diffusion signals. It provides a measurement of an effective or apparent diffusion tensor, **D**, in each voxel of an imaging volume. The DTI model assumes that the

distance travelled by water molecules in the medium vs. time follows the Einstein relationship with a probabilistic diffusion displacement profile (or "propagator") represented by an ellipsoid.

From **D** it is possible to calculate quantitative scalar and vector-valued parameters that characterize specific features of the diffusion process. These scalar quantities are rotationally invariant, i.e., independent of the coordinate system in which the MR measurement is made and the orientation of subjects in the magnet.

The most basic rotationally invariant quantities are the three principal diffusivities (or eigenvalues) of **D**, which are the principal "apparent" diffusion coefficients measured along the three intrinsic coordinate directions that constitute the local diffusion frame of reference in each voxel. Each eigenvalue is associated with a principal direction (eigenvector). In each voxel, the eigenvalues can be sorted in order of decreasing magnitude (highest diffusivity, intermediate diffusivity, and lowest diffusivity). In anisotropic tissues consisting of ordered parallel bundles, the largest eigenvalue represents the diffusion coefficient along the direction parallel to the fibers, while the other two eigenvalues represent the diffusion coefficients in two orthogonal directions perpendicular to the fibers. For this reason, the largest eigenvalue is also called ***"Axial Diffusivity (AD)***" and the mean of the other two eigenvalues is called "***Radial Diffusivity (RD)***". The average of the three eigenvalues is called "***Mean Diffusivity (MD)***". Other metrics derived from the eigenvalues are indexes of diffusion anisotropy. The most popular diffusion anisotropy measure is the "***Fractional Anisotropy (FA)***", which was named as such because it measures the fraction of the magnitude of diffusion tensor that can be attributed to its anisotropic component.

In addition to these classical diffusion tensor derived quantities, we used a simple dual compartment model to extract the "***Cerebrospinal Fluid Signal Fraction (CSF-SF)***" and the "***Parenchymal Mean Diffusivity (Par-MD)***". Par-MD is informative of the diffusivity of water molecules that are in fast exchange with the tissue's protein constituents end it should be selectively sensitive to interstitial edema.

7.2.3 Outcome Metrics Computed from the Mean Apparent Propagator

The diffusion process within a single voxel in the tissue can be more complicated than a Gaussian function as modelled by the diffusion tensor due to the hetereogeneity of the voxel in terms of different tissue types, barriers between different compartments, permeability etc... The Mean Apparent Propagator (MAP-MRI) extends DTI to generate a true and proper propagator, i.e. a spatial probability distribution function that indicates the likelihood a water

molecule will end up at a given position for a specified diffusion time. Because of its non-parametric nature and lack of any assumptions on this distribution, MAPMRI can quantify the non-Gaussian character of the diffusion process, therefore, more accurately characterizes diffusion directionality and anisotropy.

Similar to DTI, MAPMRI provides several new quantitative parameters, or MRI "stains," derived from the entire water displacement profile that captures distinct novel features about nervous tissue microstructure:

Propagator Anisotropy (PA): A more accurate metric of anisotropy can be computed with MAPMRI by relating the entire three-dimensional apparent propagator to a measure of anisotropy. This is accomplished by computing the dissimilarity of a MAPMRI propagator from its fully isotropic counterpart. Unlike its DTI counterpart FA, PA does not underestimate anisotropy in regions containing crossing-fibers and provides a more accurate description of fiber populations' angular variability within the tissue.

Non-Gaussianity (NG): MAPMRI propagators are described as a summation of weighted Hermite basis functions, where the first term is the diffusion tensor, i.e a Gaussian function. Therefore, any additional terms require to describe the propagator accurately, provide information about the deviations from Gaussianity. The NG index provides a normalized magnitude of these terms. In other words, a low NG value would indicate that the diffusion signal could be well characterized by a diffusion tensor, whereas a large NG value would indicate a significant deviation from it potentially due to restriction of water diffusion by myelin in white matter. NG is also affected by tissue hetereogeneity, or non mono-exponential diffusion signals due to noise.

Return To Origin Probability (RTOP): As described above, MAPMRI, in essence, describes a probability distribution function of where a water molecule could travel to within a specified amount of time. RTOP measures the probability that the water molecule would return to its original location. For a large cell with no barriers, this probability would be quite low, whereas for a very small cell with impermeable cell membranes, this probably would be significantly larger as the molecule does not have freedom to move farther. Therefore, RTOP is inversely proportional to pore volume. The mean diffusivity metric from the DTI model is known to be directly proportional to the pore volume, therefore, RTOP is a measure of (inverse) diffusivity for more complex diffusion profiles. The analytical mapping between the two measures have been described for Gaussian diffusion in the literature.

Return To Axis Probability (RTAP): Similar to RTOP, RTAP describes the probability of the water molecule to return to the axis of principal diffusion direction (primary eigenvector) from its initial position. This quantity is simply the reciprocal of the mean cross-sectional area of the cell, therefore related to the radial diffusivity in a Gaussian displacement profile.
Return To Plane Probability (RTPP): For diffusion taking place within coherently oriented cylinders, as in coherent fiber bundles, RTPP is equal to the reciprocal of the mean length of the cylinders, therefore inversely proportional to axial diffusivity.

In white matter we computed all the outcome metrics listed above, in gray matter the outcome metrics studied where only those meaningful for isotropic tissue, ie. MD, RTOP, Par-MD, CSF_FS. The methodology for the computation of the diffusion MRI outcome metrics is described in eAppendix2.

eAppendix7.3 Resting State fMRI Outcome Metrics

The resting state fMRI (RS-fMRI) technique can extract functional information from BOLD signals as a result of metabolic activity in the brain during resting state conditions.[67] Since the technique's first demonstration by Biswal et al., in 1995,[53] several studies emerged that postulated that brain regions which are sparsely located, share information due to the synchronization of their BOLD signals across time[46-52]. This is more popularly defined as functional connectivity (FC).[49,51,52,68-70] The intrinsic functional activity between brain regions can be assessed by evaluating FC measures using RS-fMRI to reliably extract the baseline functional networks (FNs) of the human brain. An FN comprises multiple functionally connected brain regions with correlated activation patterns across time. It can be an important outcome measure because FC can be altered in networks due to functional changes originating from several symptoms reported in AHIs.

We computed FC 'within' large-scale resting state networks using the same atlas ROIs used by Verma et al.[11] The term 'within' here means that the FC is estimated by the averaging the correlation values from the unique ROI-combinations of 'only' the ROIs comprising the specific network. The atlas used and the method incorporated to estimate 'within-network' FC are described in detail in eAppendix 4.3 and 4.4.

eTable 1. The regions of interest (ROIs) from Freesurfer combined or excluded from the analysis

Regions combined for the analysis	
Individual regions in FREESURFER	**Combined Freesurfer region for the study**
CC:_posterior, mid_Posterior, Central, Mid_Anterior	Corpus Callosum mid Sagittal Plane
Brain-Stem Left and right Cerebellum-White-Matter, CerebralWhiteMatter, TotalGrayVol	Total Parenchyma
Eliminated regions from the analysis	
CSF, Left-choroid-plexus, Right-choroid-plexus, Left and Right_vessel, Left and Right Hemisphere Surface_Hole, Optic-Chiasm, Supratentorial_volume_not_vent, supratentorial_volume_notventvox	

The table details the regions that were combined to generate a single ROI and few ROIs that were not included in the analysis. This approach was adopted to eliminate increased variabilities added to the analysis due to the small size of the ROIs.

eTable 2. **Visit information for four-healthy volunteers participating in the reproducibility analysis of the current study**

Participant	Visit2-Visit1 in days	Visit3-Visit2 in days	Visit4-Visit3 in days	Visit5-Visit4 in days
HV1	1	4	1	9
HV2	1	4	2	7
HV3	6	1	1	5
HV4	2	1	4	2

The table shows the time duration in days between each visit for the four healthy volunteers (HV) scanned as part of the NIH DTI and structural reproducibility study.

eTable 3. The average coefficient of variation (CV) calculated from averaging CV across four subjects for the interscan reproducibility analysis of volumetrics

Measure: Volume	**Average Coefficient of Variation (CV) % from 4 Healthy Volunteers**
Total Intra Cranial Volume MONSTR	0.3
Total Parenchyma	0.5
Total Gray Matter	0.6
Cerebral White Matter	0.6
White Matter Hypointensities	10.6
Cortical Gray	0.7
Sub Cortical Gray	1.0
Brainstem	1.2
Corpus Callosum midsagittal plane	2.4
Thalamus Right	2.6
Thalamus Left	3.2
Ventral Diencephalon Right	3.0
Ventral Diencephalon Left	2.6
Caudate Right	1.1
Caudate Left	1.3
Putamen Right	1.7
Putamen Left	1.1
Pallidum Right	3.9
Pallidum Left	3.0
Hippocampus Right	3.5
Hippocampus Left	1.4
Amygdala Right	3.0
Amygdala Left	3.7
Accumbens Area Right	5.8
Accumbens Area Left	6.2
Lateral Ventricle Right	2.3
Lateral Ventricle Left	2.6
Inferior Lateral Ventricle Right	9.5
Inferior Lateral Ventricle Left	4.3
3rd Ventricle	2.1
4th Ventricle	2.8
Cerebellum White Matter Right	2.1
Cerebellum White Matter Left	2.6
Cerebellum Cortex Right	0.7
Cerebellum Cortex Left	0.7
Median (Q1, Q3) CV across all regions	**2.4 (1.1, 3.2)**

The table shows the coefficient of variation (CV) %, averaged across four healthy volunteers across all regions of interest (ROIs) used in the structural volumetric analysis. It can be observed that the CV was < 5% CV for 31 of 35 ROIs. White matter hyperintensities, Accumbens left and right, inferior lateral ventricle right showed higher CV.

eTable 4. The parameter estimates of the linear regression between the Siemens Biograph mMR and the Siemens Prisma scanner

Regions of Interest (ROIs)	Siemens Biograph mMR Volume (mm^3) median (Q1, Q3)	Siemens Prisma Volume (mm^3) median (Q1, Q3)	Slope Coefficient [95% CIs]	Intercept Coefficient [95% CIs]	Correlation, *r*	Variance Explained, R^2
Total Intra Cranial Volume MONSTR	1.4e+06 (1.3e+06, 1.5e+06)	1.4e+06 (1.3e+06, 1.6e+06)	1 [0.99, 1.01]	15188.5 [-5159, 35536]	1.00	0.99
Total Parenchyma	1.2e+06 (1.1e+06, 1.3e+06)	1.2e+06 (1.1e+06, 1.3e+06)	0.97 [0.96, 0.99]	36777.9 [17615.8, 55940.1]	1.00	0.99
Total Gray Matter	6.4e+05 (6.0e+05, 6.9e+05)	6.6e+05 (6.2e+05, 7.0e+05)	0.98 [0.96, 1.01]	23219.7 [5324.9, 41114.5]	0.99	0.98
Cerebral White Matter	4.8e+05 (4.4e+05, 5.2e+05)	4.7e+05 (4.3e+05, 5.1e+05)	0.93 [0.91, 0.95]	25766.5 [14135.3, 37397.7]	0.99	0.98
White Matter Hypointensities	9.0e+02 (7.2e+02, 1.2e+03)	7.9e+02 (5.7e+02, 1.0e+03)	0.91 [0.81, 1.01]	-36.1 [-145.1, 72.8]	0.87	0.75
Cortical Gray	4.7e+05 (4.4e+05, 5.1e+05)	4.9e+05 (4.5e+05, 5.2e+05)	1 [0.97, 1.03]	16403 [892.9, 31913.1]	0.99	0.97
Sub Cortical Gray	6.2e+04 (5.7e+04, 6.6e+04)	6.0e+04 (5.5e+04, 6.3e+04)	0.94 [0.9, 0.98]	1169.4 [-1360.7, 3699.4]	0.97	0.95
Brainstem	2.3e+04 (2.1e+04, 2.4e+04)	2.2e+04 (2.0e+04, 2.4e+04)	0.96 [0.91, 1.01]	667.8 [-484.8, 1820.3]	0.96	0.93
Corpus Callosum Midsagittal Plane	3.7e+03 (3.3e+03, 4.0e+03)	3.7e+03 (3.3e+03, 4.0e+03)	0.93 [0.87, 0.98]	259.3 [47.1, 471.6]	0.95	0.91
Thalamus Right	8.0e+03 (7.4e+03, 8.6e+03)	7.4e+03 (6.8e+03, 7.9e+03)	0.87 [0.79, 0.95]	488.7 [-182.9, 1160.3]	0.89	0.80
Thalamus Left	8.2e+03 (7.5e+03, 8.9e+03)	7.7e+03 (7.1e+03, 8.4e+03)	0.83 [0.75, 0.92]	952.6 [241.2, 1663.9]	0.88	0.77

Regions of Interest (ROIs)	Siemens Biograph mMR median (Q1, Q3)	Siemens Prisma median (Q1, Q3)	Slope Coefficient [95% CIs]	Intercept Coefficient [95% CIs]	Correlation, *r*	Variance Explained, R^2
Ventral Diencephalon Right	4.1e+03 (3.9e+03, 4.4e+03)	4.2e+03 (3.8e+03, 4.5e+03)	0.83 [0.74, 0.91]	728.1 [364.9, 1091.3]	0.88	0.77
Ventral Diencephalon Left	4.2e+03 (3.9e+03, 4.4e+03)	4.2e+03 (3.9e+03, 4.5e+03)	0.85 [0.77, 0.93]	609.6 [265.5, 953.7]	0.90	0.80
Caudate Right	3.9e+03 (3.4e+03, 4.2e+03)	3.9e+03 (3.5e+03, 4.2e+03)	0.94 [0.88, 1]	288.2 [64.4, 512]	0.95	0.91
Caudate Left	3.7e+03 (3.3e+03, 4.0e+03)	3.7e+03 (3.4e+03, 4.1e+03)	0.94 [0.9, 0.99]	246.7 [85.3, 408.1]	0.97	0.95
Putamen Right	5.3e+03 (4.8e+03, 5.7e+03)	5.1e+03 (4.7e+03, 5.5e+03)	0.93 [0.88, 0.97]	238.1 [23.9, 452.3]	0.97	0.95
Putamen Left	5.3e+03 (4.9e+03, 5.8e+03)	5.0e+03 (4.6e+03, 5.5e+03)	0.93 [0.88, 0.98]	97.8 [-170.4, 365.9]	0.96	0.93
Pallidum Right	2.1e+03 (1.9e+03, 2.2e+03)	2.0e+03 (1.8e+03, 2.2e+03)	0.77 [0.68, 0.87]	417.7 [215.8, 619.5]	0.83	0.70
Pallidum Left	2.2e+03 (2.0e+03, 2.3e+03)	2.1e+03 (1.9e+03, 2.3e+03)	0.76 [0.68, 0.84]	466.6 [284.6, 648.7]	0.87	0.75
Hippocampus Right	4.4e+03 (4.2e+03, 4.7e+03)	4.1e+03 (3.9e+03, 4.5e+03)	0.91 [0.85, 0.97]	161.2 [-115.4, 437.8]	0.94	0.89
Hippocampus Left	4.3e+03 (4.0e+03, 4.5e+03)	4.0e+03 (3.8e+03, 4.3e+03)	0.87 [0.8, 0.93]	311.5 [29.9, 593.2]	0.93	0.86
Amygdala Right	1.8e+03 (1.7e+03, 1.9e+03)	1.7e+03 (1.6e+03, 1.9e+03)	0.96 [0.89, 1.03]	-8 [-140.1, 124.2]	0.93	0.87

Regions of Interest (ROIs)	Siemens Biograph mMR median (Q1, Q3)	Siemens Prisma median (Q1, Q3)	Slope Coefficient [95% CIs]	Intercept Coefficient [95% CIs]	Correlation, *r*	Variance Explained, R^2
Amygdala Left	1.7e+03 (1.5e+03, 1.9e+03)	1.6e+03 (1.5e+03, 1.8e+03)	0.83 [0.73, 0.94]	176.6 [-0.9, 354.1]	0.84	0.71
Accumbens Area Right	5.6e+02 (5.1e+02, 6.1e+02)	5.0e+02 (4.5e+02, 5.6e+02)	0.73 [0.62, 0.84]	92.6 [30.5, 154.7]	0.79	0.62
Accumbens Area Left	5.4e+02 (4.9e+02, 6.2e+02)	4.5e+02 (3.9e+02, 5.1e+02)	0.54 [0.41, 0.66]	154.7 [85.5, 224]	0.64	0.41
Lateral Ventricle Right	7.0e+03 (5.1e+03, 9.8e+03)	7.0e+03 (5.1e+03, 9.7e+03)	1 [0.99, 1.01]	8.3 [-103.2, 119.8]	1.00	1.00
Lateral Ventricle Left	7.5e+03 (6.0e+03, 1.0e+04)	7.6e+03 (6.1e+03, 1.0e+04)	1 [0.98, 1.01]	21.2 [-111.1, 153.5]	1.00	0.99
Inferior Lateral Ventricle Right	3.4e+02 (2.4e+02, 4.3e+02)	3.3e+02 (2.5e+02, 4.4e+02)	0.85 [0.79, 0.91]	55.9 [32.6, 79.2]	0.94	0.88
Inferior Lateral Ventricle Left	3.2e+02 (2.5e+02, 4.1e+02)	3.3e+02 (2.6e+02, 4.3e+02)	0.84 [0.74, 0.93]	62.3 [27.3, 97.2]	0.85	0.73
Third Ventricle	9.1e+02 (7.6e+02, 1.1e+03)	9.3e+02 (8.0e+02, 1.1e+03)	0.97 [0.93, 1.01]	47.1 [12.1, 82.2]	0.98	0.96
Fourth Ventricle	1.8e+03 (1.5e+03, 2.3e+03)	1.9e+03 (1.6e+03, 2.3e+03)	0.97 [0.93, 1.02]	124.9 [27.8, 221.9]	0.97	0.93
Cerebellum White Matter Right	1.5e+04 (1.4e+04, 1.6e+04)	1.5e+04 (1.3e+04, 1.6e+04)	0.95 [0.78, 1.12]	526.3 [-2012.5, 3065]	0.73	0.53
Cerebellum White Matter Left	1.6e+04 (1.5e+04, 1.7e+04)	1.5e+04 (1.4e+04, 1.6e+04)	0.86 [0.72, 0.99]	1725 [-386.2, 3836.3]	0.78	0.60

Regions of Interest (ROIs)	Siemens Biograph mMR median (Q1, Q3)	Siemens Prisma median (Q1, Q3)	Slope Coefficient [95% CIs]	Intercept Coefficient [95% CIs]	Correlation, *r*	Variance Explained, R^2
Cerebellum Cortex Right	5.5e+04 (5.1e+04, 5.9e+04)	5.4e+04 (5.1e+04, 5.9e+04)	0.91 [0.87, 0.95]	4107.8 [1878.5, 6337]	0.97	0.95
Cerebellum Cortex Left	5.4e+04 (5.1e+04, 5.8e+04)	5.4e+04 (5.1e+04, 5.8e+04)	0.89 [0.85, 0.93]	5542.2 [3229.3, 7855.1]	0.97	0.94

The median volume (in mm^3) and the corresponding quartiles for each ROI are shown for both scanners (Columns 2, 3). The slope and intercept coefficients and their 95% confidence intervals (CIs) for each ROI are provided, (Columns 4, 5) derived from the linear relationship between the two scanners. The correlation (***r***) between the measurements from the two scanners and the $\mathbf{R^2}$ values are presented in the last two columns. 30 out of the 35 ROIs (~ 86%) presented in the table demonstrated an $\mathbf{R}^2 > 0.70$. Abbreviations: $\mathbf{R^2}$ = the coefficient of determination that estimates the % of variance in one scanner explained by the other scanner.

eTable 5. The diffusion gradients that are input to the scanner

Anterior-Posterior/Posterior-Anterior (AP/PA) maximum b=1100 s/mm^2				**Anterior-Posterior/Posterior-Anterior (AP/PA) maximum b=2500 s/mm^2**				**Right-Left/Left-Right (RL/LR) maximum b=1100 s/mm^2**				**Right-Left/Left-Right (RL/LR) maximum b=2500 s/mm^2**			
X	Y	Z	bval	X	Y	Z	bval	X	Y	Z	bval	X	Y	Z	bval
0.000	0.000	0.000	0	0.000	0.000	0.000	0	0.000	0.000	0.000	0	0.000	0.000	0.000	0
0.062	0.131	0.156	50	0.041	0.087	0.104	50	-0.135	-0.146	0.079	50	-0.089	-0.096	0.052	50
-0.170	0.098	0.083	50	-0.113	0.065	0.055	50	0.098	-0.113	0.152	50	0.065	-0.074	0.101	50
0.000	0.000	0.522	300	-0.139	-0.888	0.438	2500	0.522	0.000	0.000	300	0.100	0.082	0.992	2500
-0.290	-0.323	0.290	300	-0.605	-0.059	0.794	2500	0.000	0.522	0.000	300	0.872	0.438	0.221	2500
0.342	0.939	0.023	1100	-0.831	0.553	0.061	2500	0.485	0.826	0.288	1100	0.971	-0.220	0.096	2500
-0.158	0.915	0.372	1100	0.754	-0.027	0.656	2500	-0.566	0.821	0.076	1100	-0.255	-0.637	0.727	2500
0.748	0.628	0.213	1100	0.744	0.553	0.376	2500	0.852	0.046	0.522	1100	-0.802	-0.222	0.555	2500
-0.762	-0.068	0.645	1100	-0.899	-0.395	0.188	2500	0.556	-0.517	0.651	1100	0.741	-0.604	0.294	2500
0.080	0.270	0.959	1100	0.139	-0.511	0.848	2500	0.020	-0.831	0.555	1100	-0.511	0.688	0.515	2500
0.253	-0.595	0.763	1100	-0.066	0.399	0.915	2500	-0.923	0.033	0.385	1100	0.313	0.929	0.198	2500
0.941	-0.172	0.292	1100	-0.937	-0.032	-0.349	2500	0.158	0.554	0.817	1100	0.639	0.732	-0.236	2500
-0.399	-0.534	0.746	1100	-0.032	-0.887	-0.461	2500	-0.194	-0.340	0.920	1100	-0.421	0.652	-0.631	2500
-0.318	-0.760	-0.567	1100	-0.515	-0.360	-0.778	2500	0.245	-0.266	-0.932	1100	-0.455	-0.009	-0.891	2500
0.309	0.035	-0.950	1100	0.637	-0.412	-0.652	2500	-0.030	-0.999	-0.039	1100	0.875	-0.206	-0.439	2500
-0.611	0.002	-0.792	1100	0.109	0.184	-0.977	2500	-0.302	0.042	-0.952	1100	0.358	0.192	-0.914	2500
-0.747	0.654	-0.122	1100	-0.600	0.364	-0.712	2500	0.600	0.647	-0.471	1100	-0.859	0.221	-0.462	2500
0.462	0.843	-0.276	1100	-0.447	0.811	-0.379	2500	0.331	-0.710	-0.622	1100	0.034	-0.661	-0.749	2500
0.648	-0.563	-0.513	1100	0.552	0.391	-0.736	2500	-0.606	0.686	-0.402	1100	0.090	-0.978	-0.191	2500
-0.145	0.955	-0.257	1100	-	-	-	-	-0.914	0.405	-0.038	1100	-	-	-	-
0.974	0.185	-0.134	1100	-	-	-	-	-0.848	-0.359	-0.390	1100	-	-	-	-

eTable 6. The coefficient of variation (CV) across repeated scans for diffusion MRI metrics in white matter and gray matter regions

White Matter Regions of Interest	Coefficient of Variation (CV) %, for Diffusion MRI Metrics										
	FA	MD	Par-MD	CSF-SF	AD	RD	NG	PA	RTAP	RTOP	RTPP
Middle_cerebellar_peduncle	0.6	0.6	1.3	6.1	0.4	1.2	0.8	0.1	1.1	1.2	0.3
Pontine_crossing_tract_a_part_of_MCP	1.5	0.7	2.5	11.6	0.6	1.3	1.7	0.4	1.0	1.1	0.4
Genu_of_corpus_callosum	0.3	0.5	1.2	9.2	0.5	1.1	1.7	0.1	1.9	1.8	0.2
Body_of_corpus_callosum	0.5	0.4	1.0	3.9	0.5	1.0	1.2	0.1	1.3	1.4	0.2
Splenium_of_corpus_callosum	0.3	0.6	1.0	7.0	0.4	1.2	0.9	0.1	1.2	1.4	0.2
Fornix	0.7	0.9	3.8	8.0	1.1	1.0	2.9	0.5	2.8	2.7	0.7
Corticospinal_tract_R	0.6	0.8	4.7	9.2	0.9	1.1	1.4	0.4	2.2	2.3	0.6
Corticospinal_tract_L	0.9	0.7	3.4	6.3	0.6	1.3	1.0	0.2	1.6	1.4	0.4
Medial_Lemniscus_R	0.5	0.5	1.7	5.5	0.3	0.9	2.1	0.4	2.3	2.1	0.3
Medial_Lemniscus_L	0.8	0.8	2.7	9.6	0.7	1.5	2.2	0.3	1.6	1.4	0.4
Inferior_cerebellar_peduncle_R	0.8	0.7	2.7	7.0	0.5	1.1	2.0	0.5	1.6	1.5	0.3
Inferior_cerebellar_peduncle_L	0.7	0.7	2.2	5.5	0.7	0.9	2.1	0.6	1.3	1.3	0.4
Superior_cerebellar_peduncle_R	1.2	0.9	3.5	5.8	0.6	1.8	1.5	0.2	1.8	1.7	0.3
Superior_cerebellar_peduncle_L	0.7	0.9	3.0	5.6	0.8	1.4	1.8	0.3	1.7	1.5	0.4
Cerebral_peduncle_R	0.6	0.8	3.7	13.8	0.6	1.8	1.4	0.1	2.2	2.3	0.3
Cerebral_peduncle_L	0.5	0.6	3.0	6.3	0.6	1.3	1.0	0.1	1.2	1.4	0.3
Anterior_Limb_of_internal_capsule_R	0.6	0.6	3.0	11.7	0.4	1.1	1.2	0.2	1.5	1.6	0.3

White Matter Regions of Interest	FA	MD	Par-MD	CSF-SF	AD	RD	NG	PA	RTAP	RTOP	RTPP
Anterior_Limb_of_internal_capsule_L	0.7	0.7	2.1	7.5	0.7	1.0	1.5	0.2	1.6	1.6	0.4
Posterior_Limb_of_internal_capsule_R	0.4	0.6	2.3	7.6	0.5	1.0	1.0	0.1	1.2	1.3	0.3
Posterior_Limb_of_internal_capsule_L	0.4	0.4	2.9	5.5	0.4	0.7	0.5	0.1	1.0	1.2	0.3
Retrolenticular_part_of_internal_capsule_R	0.7	0.7	1.3	6.6	0.5	1.2	1.5	0.1	0.9	1.1	0.3
Retrolenticular_part_of_internal_capsule_L	0.6	0.4	1.0	9.0	0.4	0.8	0.9	0.1	1.4	1.5	0.3
Anterior_corona_Radiata_R	0.6	0.5	1.3	6.1	0.4	0.7	1.2	0.4	1.1	1.3	0.2
Anterior_corona_Radiata_L	0.6	0.6	1.2	4.0	0.5	0.7	1.1	0.4	1.1	1.4	0.2
Superior_corona_Radiata_R	0.6	0.7	1.4	4.8	0.6	0.8	0.8	0.2	0.8	1.2	0.3
Superior_corona_Radiata_L	0.6	0.5	1.2	3.8	0.4	0.7	0.9	0.3	1.2	1.5	0.3
Posterior_corona_Radiata_R	0.7	0.8	1.4	8.0	0.6	1.0	1.2	0.2	0.8	1.2	0.4
Posterior_corona_Radiata_L	0.7	0.6	0.6	4.3	0.5	0.7	0.8	0.3	1.0	1.2	0.3
Posterior_thalamic_Radiation_include_optic_Radiation_R	0.6	0.8	1.0	8.2	0.6	1.2	1.7	0.2	1.1	1.3	0.4
Posterior_thalamic_Radiation_include_optic_Radiation_L	0.5	0.5	1.0	5.6	0.5	0.8	1.0	0.1	1.2	1.3	0.3
Sagittal_stratum_R	0.7	0.5	1.4	7.2	0.4	1.0	1.5	0.1	0.9	0.9	0.3
Sagittal_stratum_L	0.5	0.5	1.3	8.6	0.5	0.7	1.3	0.2	1.5	1.6	0.2
External_capsule_R	0.7	0.6	1.1	5.5	0.4	0.8	2.6	0.8	1.1	1.2	0.2

White Matter Regions of Interest	FA	MD	Par-MD	CSF-SF	AD	RD	NG	PA	RTAP	RTOP	RTPP
External_capsule_L	0.5	0.4	1.4	9.0	0.4	0.5	1.8	0.6	1.1	1.2	0.3
Cingulum_cingulate_gyrus_R	0.5	0.5	1.0	6.7	0.5	0.8	1.0	0.2	0.9	1.0	0.3
Cingulum_cingulate_gyrus_L	0.4	0.4	1.3	6.7	0.5	0.7	0.9	0.2	1.1	1.0	0.3
Cingulum_hippocampus_R	0.6	0.6	2.2	11.1	0.4	0.9	2.2	0.4	1.3	1.3	0.3
Cingulum_hippocampus_L	0.5	0.6	1.7	12.6	0.7	0.8	2.0	0.4	1.5	1.5	0.3
Fornix_cres_Stria_terminalis_R	0.9	0.8	2.5	9.5	0.7	1.4	2.5	0.3	2.0	1.7	0.4
Fornix_cres_Stria_terminalis_L	0.4	0.5	2.2	9.3	0.5	0.8	1.9	0.2	1.7	1.8	0.4
Superior_Longitudinal_fasciculus_R	0.8	0.7	1.4	9.7	0.5	1.1	1.0	0.1	1.0	1.2	0.3
Superior_Longitudinal_fasciculus_L	0.8	0.6	1.1	6.1	0.6	0.9	0.8	0.1	1.1	1.3	0.4
Superior_fronto-occipital_fasciculus_R	1.0	0.6	3.3	11.9	0.7	1.0	2.0	0.7	1.5	1.6	0.4
Superior_fronto-occipital_fasciculus_L	1.4	0.9	2.6	12.9	0.8	1.3	2.1	0.7	1.5	1.7	0.6
Uncinate_fasciculus_R	1.1	0.7	3.0	9.8	0.4	1.4	3.3	0.6	1.7	1.6	0.4
Uncinate_fasciculus_L	1.1	0.6	2.4	13.2	0.8	1.1	3.1	0.6	2.2	2.5	0.5
Tapetum_R	0.6	0.7	3.6	9.8	0.5	1.1	1.9	0.2	1.4	1.3	0.4
Tapetum_L	1.0	0.6	1.6	8.8	0.7	1.0	1.9	0.3	1.6	1.9	0.3
Cerebellum White Matter R	0.7	0.6	1.2	4.4	0.5	0.8	1.7	0.7	0.8	0.9	0.3
Cerebellum White Matter L	0.6	0.7	0.9	3.7	0.6	0.7	1.4	0.6	0.9	1.1	0.3

White Matter Regions of Interest	**FA**	**MD**	**Par-MD**	**CSF-SF**	**AD**	**RD**	**NG**	**PA**	**RTAP**	**RTOP**	**RTPP**
Median (Q1, Q3) CV across all regions	**0.6 (0.5, 0.8)**	**0.6 (0.5, 0.7)**	**1.7 (1.2, 2.8)**	**7.4 (5.6, 9.5)**	**0.5 (0.4, 0.6)**	**1 (0.8, 1.2)**	**1.5 (1.0, 2.0)**	**0.2 (0.1, 0.4)**	**1.3 (1.1, 1.6)**	**1.4 (1.2, 1.6)**	**0.3 (0.3, 0.4)**
Gray Matter Regions of Interest	**FA**	**MD**	**Par-MD**	**CSF-SF**	**AD**	**RD**	**NG**	**PA**	**RTAP**	**RTOP**	**RTPP**
Amygdala R	1.3	0.7	1.1	3.8	0.7	0.7	3.4	2.3	0.6	0.8	0.3
Amygdala L	2.0	0.6	1.4	5.6	0.8	0.6	3.0	2.4	0.6	0.8	0.4
Caudate R	1.6	1.6	2.1	4.6	1.6	1.7	2.0	1.8	1.2	1.5	0.6
Caudate L	1.3	1.8	2.0	5.1	1.7	1.8	2.1	1.3	1.2	1.4	0.5
Putamen R	1.7	0.6	1.4	7.5	0.5	0.6	3.2	2.2	1.0	1.2	0.3
Putamen L	1.3	0.5	1.2	7.4	0.5	0.5	2.0	2.2	0.7	0.9	0.2
Thalamus Proper R	0.6	1.0	1.5	4.1	1.1	1.0	2.3	1.2	0.9	1.1	0.4
Thalamus Proper L	0.7	1.1	1.6	3.8	1.0	1.2	1.8	1.0	0.8	1.0	0.4
GM Cerebellum Exterior R	1.1	0.9	2.6	3.9	1.0	1.0	2.8	2.3	1.1	1.6	0.5
GM Cerebellum Exterior L	0.9	0.9	1.6	2.7	0.9	0.9	1.7	1.4	0.8	1.1	0.5
Cerebellar Vermal Lobules I-V	2.0	0.8	2.9	2.1	0.9	0.7	1.4	1.2	0.9	1.4	0.5
Cerebellar Vermal Lobules VI-VII	1.3	0.9	2.4	2.5	0.8	0.9	1.6	1.4	0.7	1.0	0.5

Gray Matter Regions of Interest	FA	MD	Par-MD	CSF-SF	AD	RD	NG	PA	RTAP	RTOP	RTPP
Cerebellar Vermal Lobules VIII-X	0.9	1.1	3.4	4.1	1.0	1.1	2.5	2.0	0.7	1.2	0.5
Median (Q1, Q3) CV across all GM ROIs	**1.3 (0.9, 1.7)**	**0.9 (0.7, 1.1)**	**1.6 (1.4, 2.5)**	**4.1 (3.3, 5.4)**	**0.9 (0.8, 1.1)**	**0.9 (0.7, 1.2)**	**2.1 (1.8, 2.9)**	**1.8 (1.3, 2.3)**	**0.8 (0.7, 1.1)**	**1.1 (1.0, 1.4)**	**0.5 (0.4, 0.5)**

The table shows the coefficient of variation (CV) %, calculated for white matter (WM) (top) and gray matter (GM) (bottom) regions of interest (ROIs) across repeated scans; averaged across the four healthy volunteers in the reproducibility experiment of the study. The median and the first and third quartiles (Q1, Q3) of the CV across all WM and GM ROIs per metric are also shown in bold. It can be observed from the table that the median CV across both WM and GM ROIs is low (~ ≤ 2%) for all diffusion MRI metrics, except for CSF-SF.

eTable 7. Group differences in dMRI metrics from the Bayesian analysis between control and participants with AHI

Regions of Interest (ROIs)	Diffusion MRI Metric	*Bayesian* Posterior Probability *(P+)*	Participants with AHIs vs. Controls Difference in Median (%)	Significant ROI Match with Conventional Analysis	Diffusion Metric for Significant ROI Match with Conventional Analysis
Genu of Corpus Callosum	Return to Axis Probability (RTAP)	0.000	-3.359	✓	✓
	Return to Origin Probability (RTOP)	0.003	-2.704		✓
Body of Corpus Callosum	Return to Axis Probability (RTAP)	0.000	-2.782	✓	✓
	Return to Origin Probability (RTOP)	0.013	-2.042		×
Splenium of Corpus Callosum	Return to Axis Probability (RTAP)	0.001	-2.448	×	×
Superior Longitudinal Fasciculus Right	Return to Axis Probability (RTAP)	0.017	-2.021	✓	✓
Caudate Right	Non-Gaussianity (NG)	0.014	-2.933	✓	✓

Difference in Median (%) = $\frac{Median_{AHI} - Median_{Control}}{Median_{Control}} \times 100$

The regions of interest shown here are based on posterior probability values for the *Bayesian* multilevel modeling analysis, $P^{+} < 0.05$ or $P^{+} > 0.95$, and the between group difference having a magnitude of difference in median > 2% (Column 4). ROIs and diffusion MRI metrics from this test that agree or disagree with the conventional analysis are shown in the last two columns, respectively.

eTable 8. Group differences in dMRI metrics from the Kruskal-Wallis (KW) test across controls, AHI1 and AHI2 subgroups

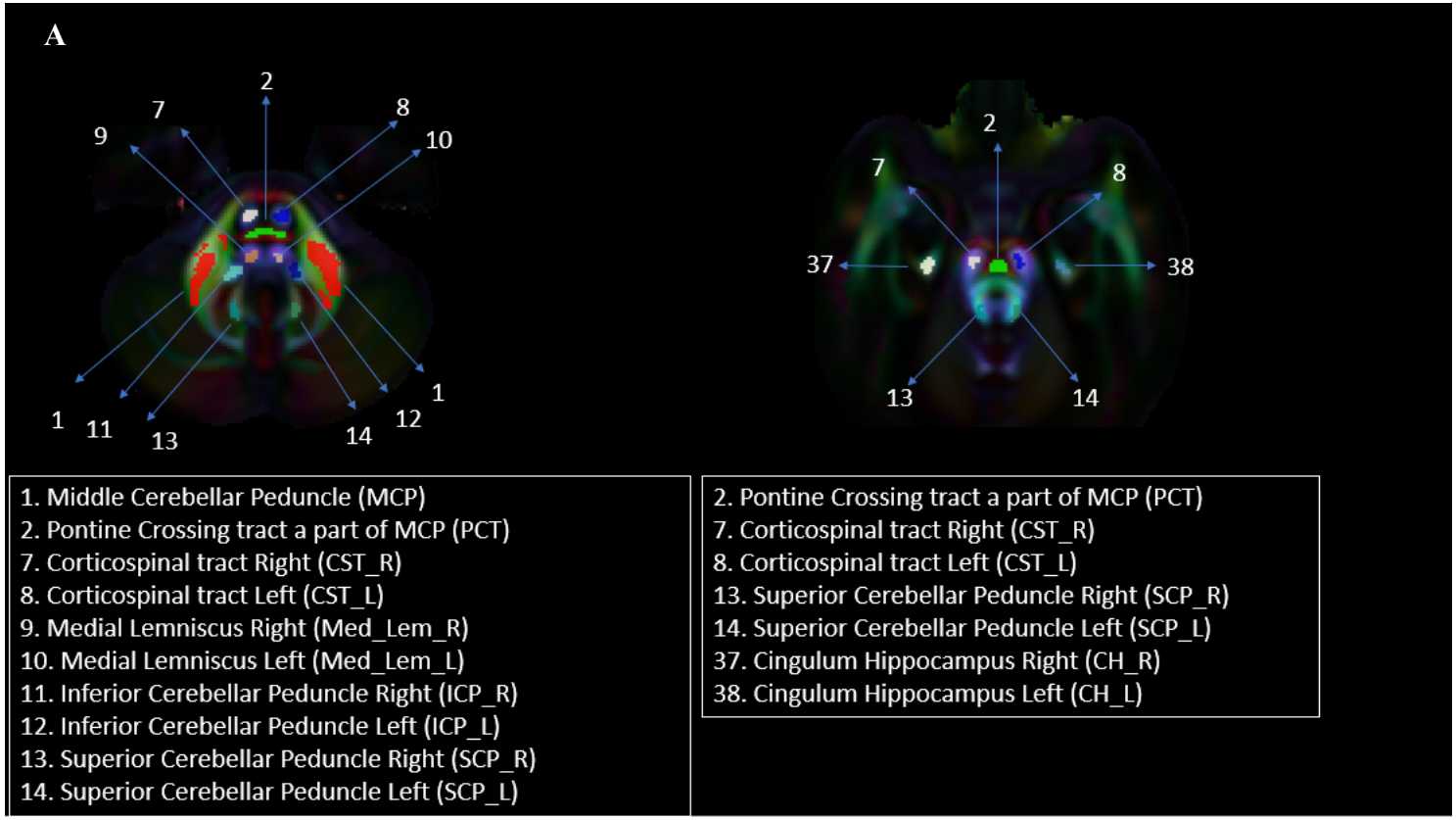
Regions of Interest (ROIs)	Diffusion MRI Metrics	Kruskal-Wallis (KW) P value unadjusted	Kruskal-Wallis (KW) P value Benjamini-Hochberg adjusted	AHI1 vs. Controls Mann-Whitney P value Benjamini-Hochberg adjusted	AHI2 vs. Controls Mann-Whitney P value Benjamini-Hochberg adjusted	AHI1 vs. Controls Difference in Median (%)	AHI2 vs. Controls Difference in Median (%)
External Capsule Right	CSF-SF	0.022	0.691	0.032	0.032	-5.35	-5.26
Superior Longitudinal Fasciculus Right	FA	0.006	0.151	0.128	0.004	-2.42	-3.76
	RD	0.017	0.546	0.466	0.019	0.88	4.33
	RTAP	0.016	0.742	0.348	0.011	-0.13	-4.89
Superior Longitudinal Fasciculus Left	FA	0.020	0.315	0.327	0.028	-0.17	-3.74
	RD	0.033	0.689	0.386	0.036	0.50	4.10
Superior Fronto-occipital Fasciculus Right	FA	0.038	0.378	0.049	0.915	4.49	-0.99
	RD	0.045	0.715	0.116	0.379	-1.30	2.20
Superior Fronto-occipital Fasciculus Left	FA	0.005	0.151	0.020	0.407	5.65	-1.37
	MD	0.020	0.629	0.763	0.025	-0.11	1.97
	RD	0.013	0.546	0.110	0.161	-2.06	4.50
Pontine Crossing Fibers	AD	0.016	0.330	0.024	0.855	1.63	-0.49
Inferior Cerebellar Peduncle Right	AD	0.001	**0.025**	0.002	0.643	1.51	-0.67
	FA	0.025	0.315	0.049	0.549	1.83	-0.52

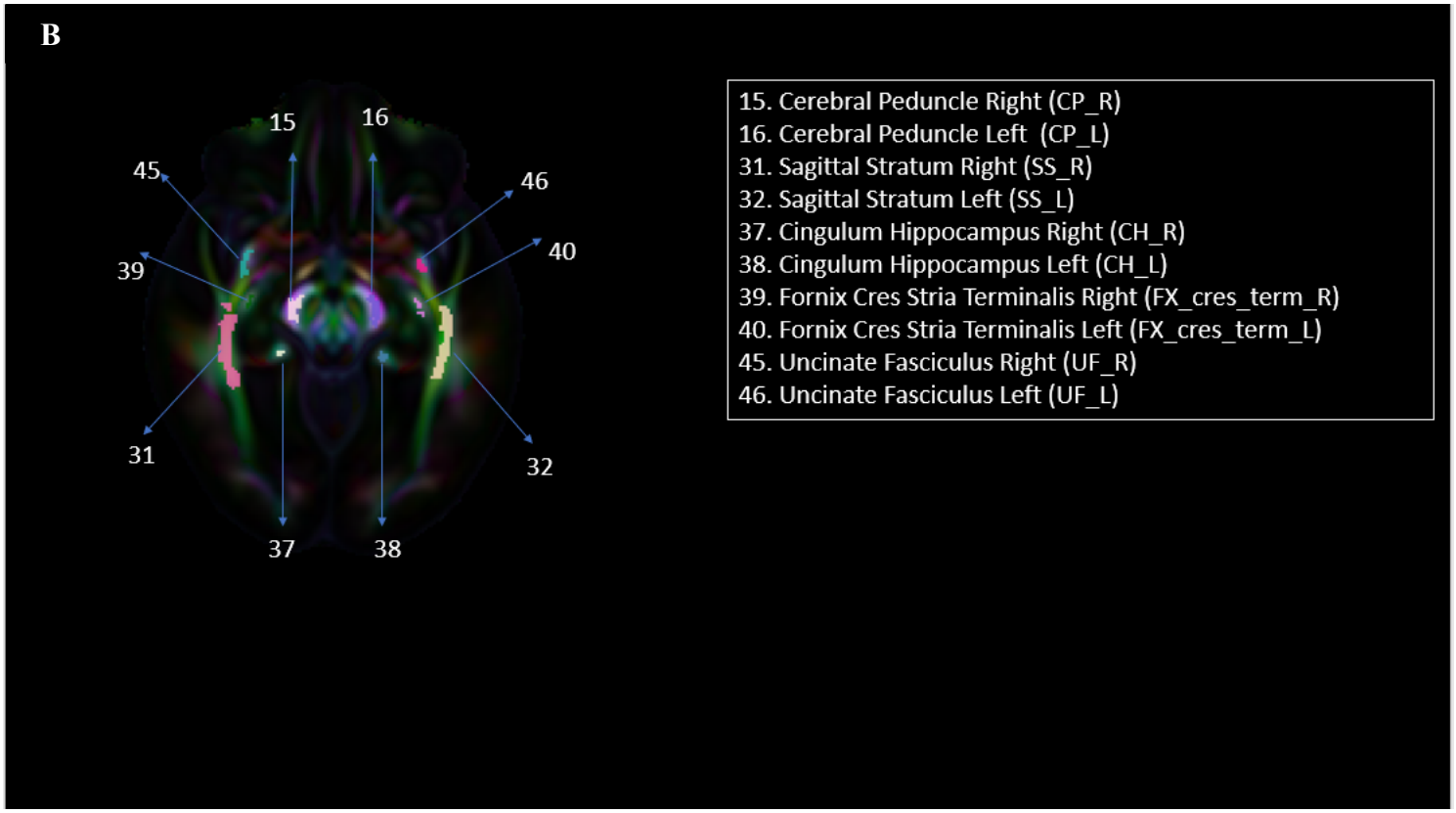
Regions of Interest (ROIs)	Diffusion MRI Metrics	Kruskal-Wallis (KW) P value unadjusted	Kruskal-Wallis (KW) P value Benjamini-Hochberg adjusted	AHI1 vs. Controls Mann-Whitney P value Benjamini-Hochberg adjusted	AHI2 vs. Controls Mann-Whitney P value Benjamini-Hochberg adjusted	AHI1 vs. Controls Difference in Median (%)	AHI2 vs. Controls Difference in Median (%)
Inferior Cerebellar Peduncle Left	AD	0.001	**0.025**	0.001	0.909	2.49	0.03
	CSF-SF	0.008	0.496	0.013	0.555	8.39	2.25
	MD	0.005	0.342	0.006	0.398	3.26	1.62
Caudate Right	NG	0.001	**0.037**	0.000	0.025	-5.47	-3.71

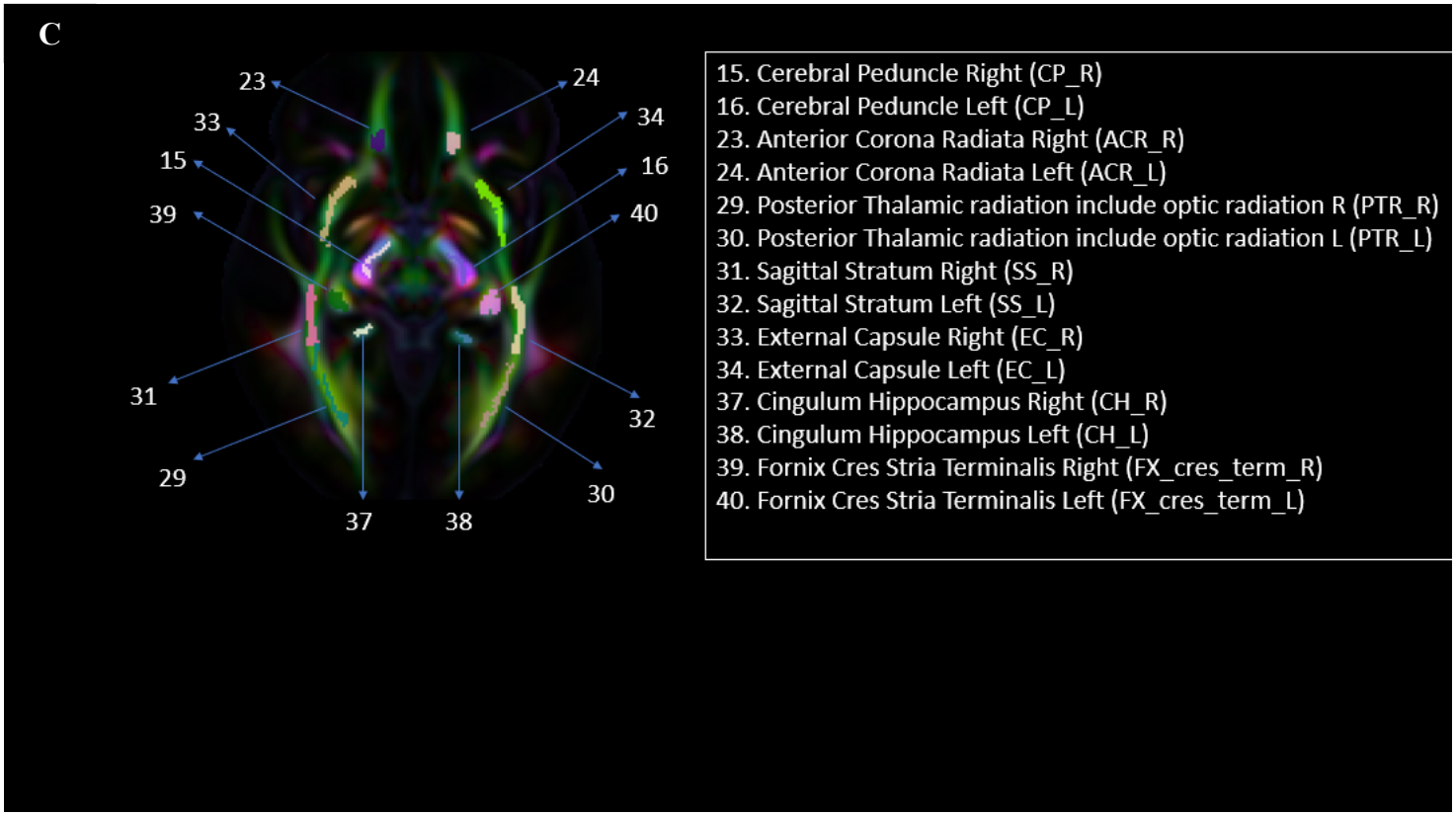
Difference in Median (%) = $\frac{Median_{AHI} - Median_{Control}}{Median_{Control}} \times 100$

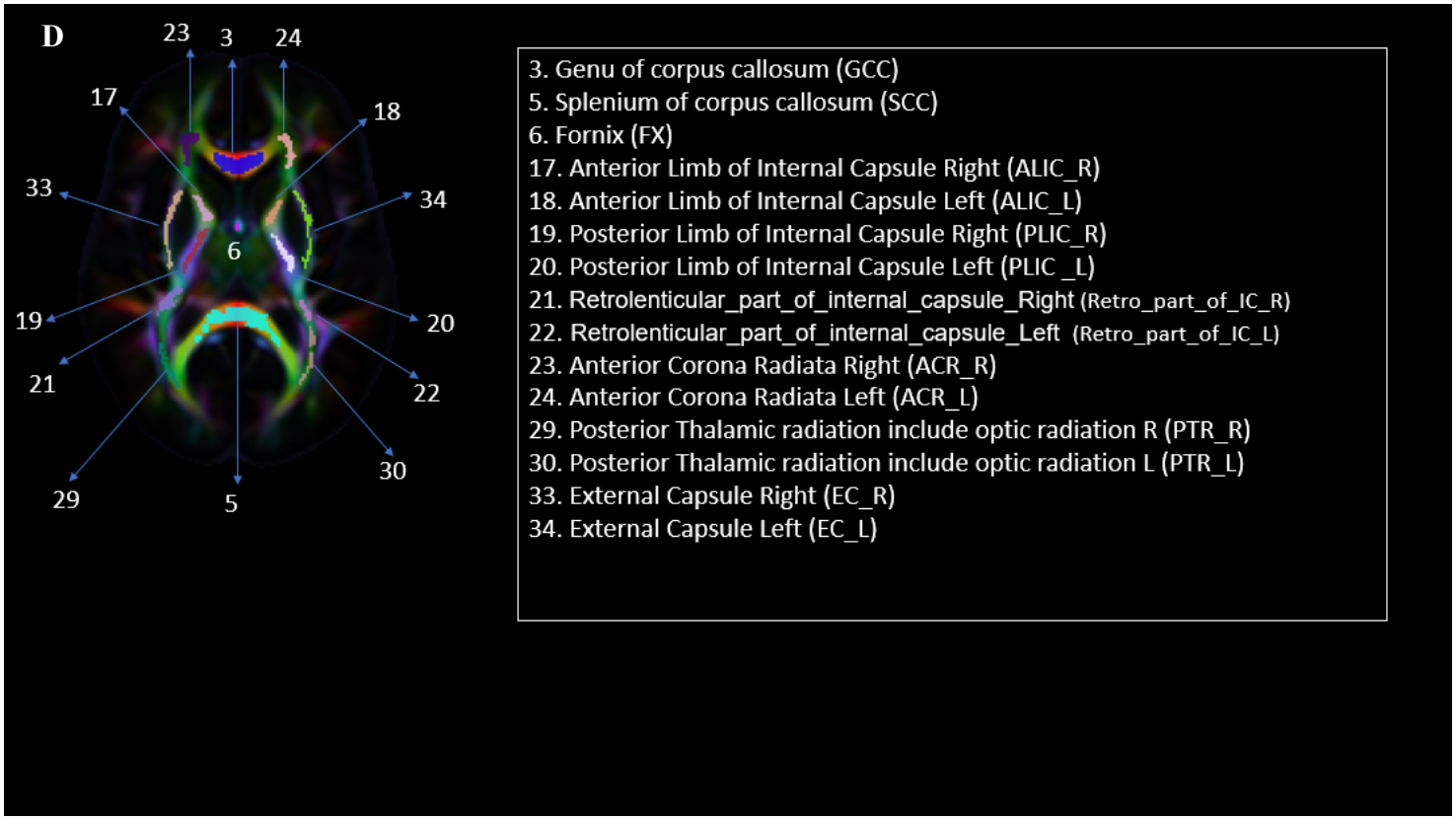
The regions of interest (ROIs) shown here are based on an unadjusted $P_{KW} < 0.05$, and if at least one pair-wise between group difference from the Mann-Whitney *U* test showed a difference in median > 2%. The Benjamini-Hochberg (BH) corrected P values for the KW tests are provided on the fourth column. Any P values that survive the BH adjustment are shown in bold. All P values shown from the pair-wise group comparisons using the Mann-Whitney *U* test are shown in columns 5-7 and they are BH corrected. Abbreviations: CSF-SF = cerebrospinal fluid signal fraction, FA = fractional anisotropy, RD = radial diffusivity, RTAP = return to axis probability, MD = mean diffusivity, AD = axial diffusivity and NG = non-Gaussianity

eFigures

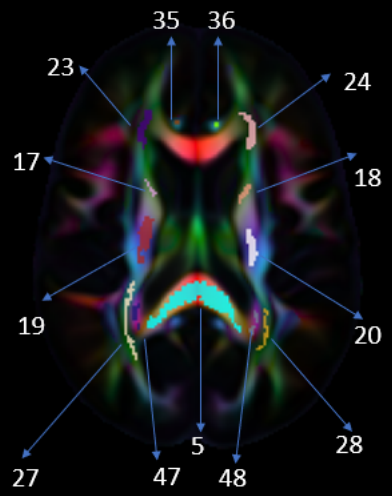




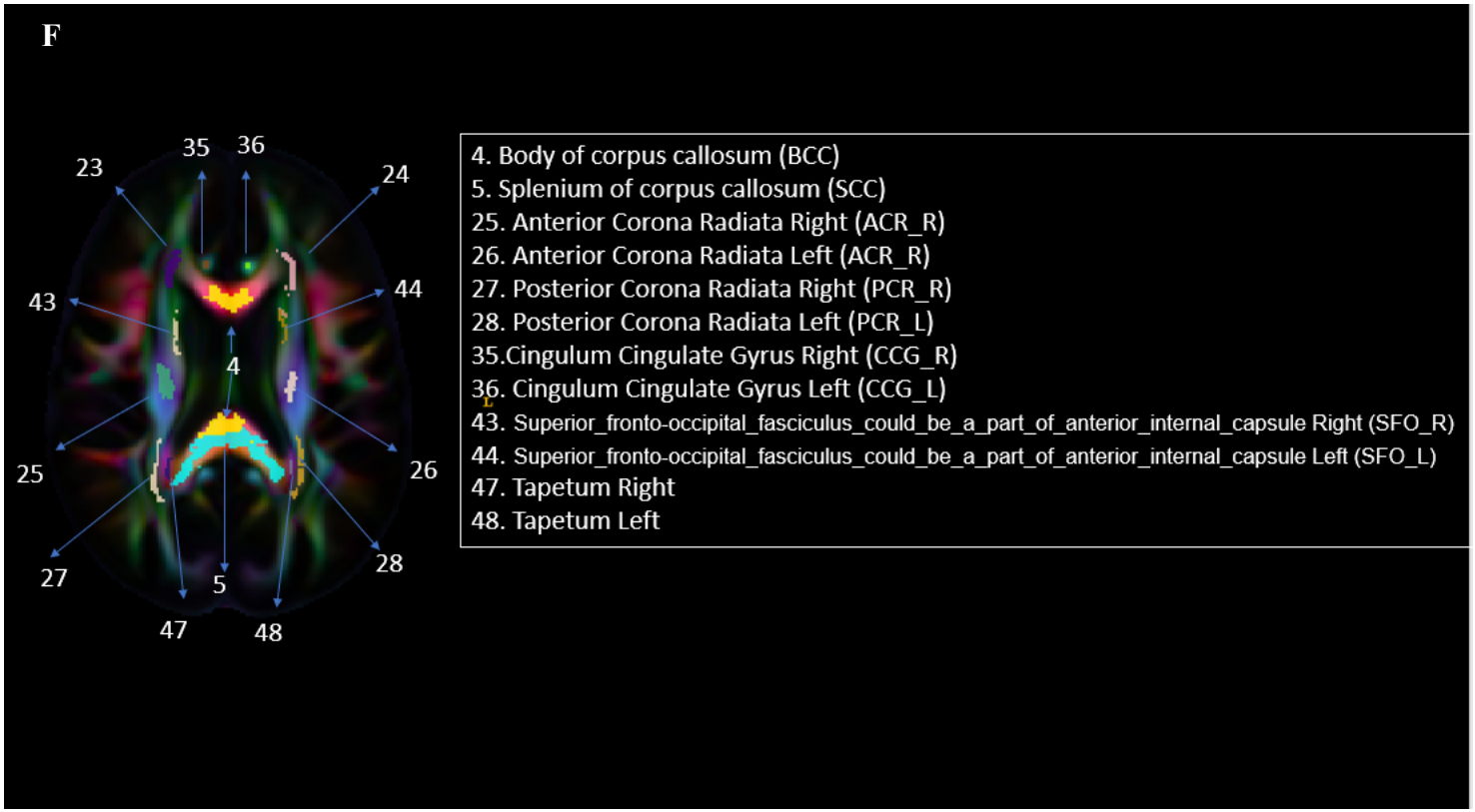


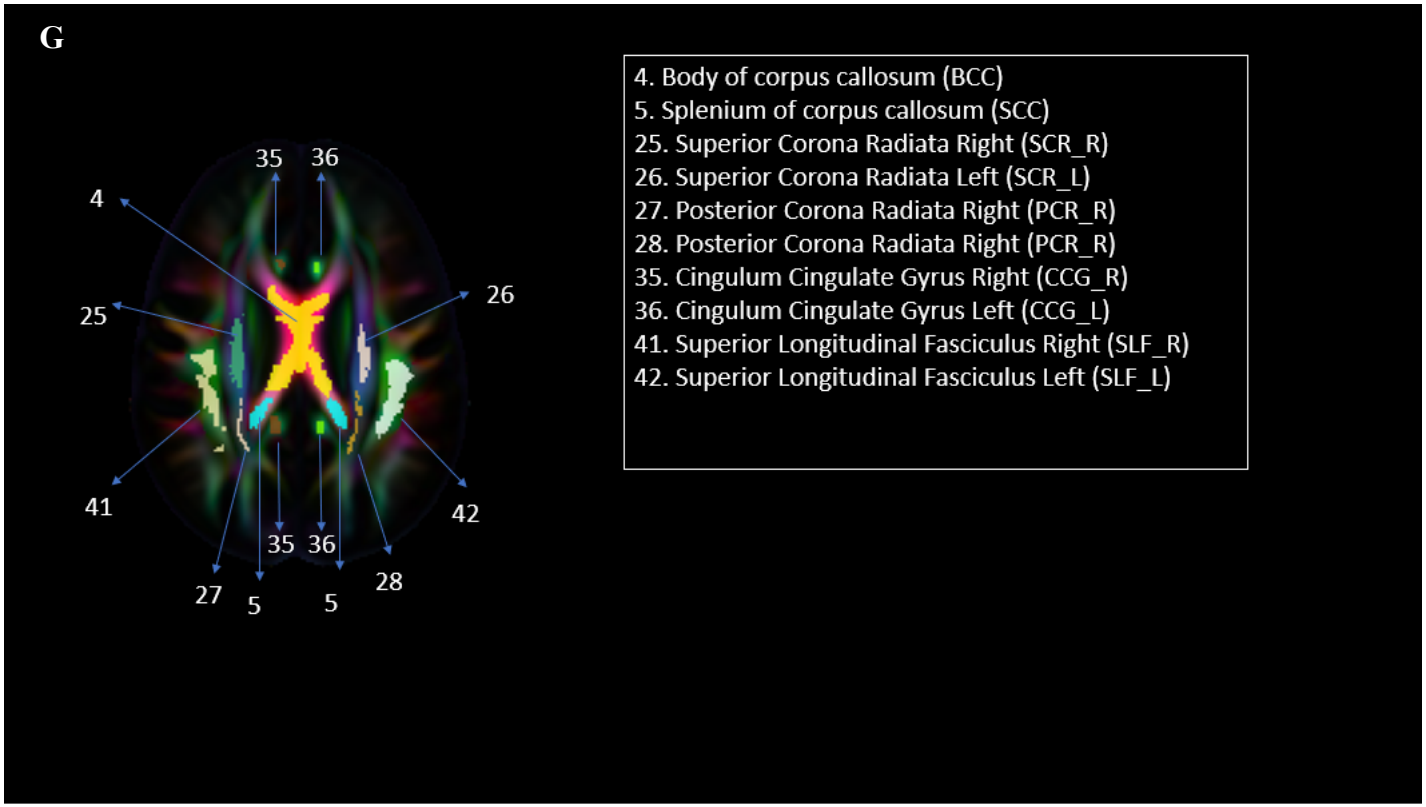


E

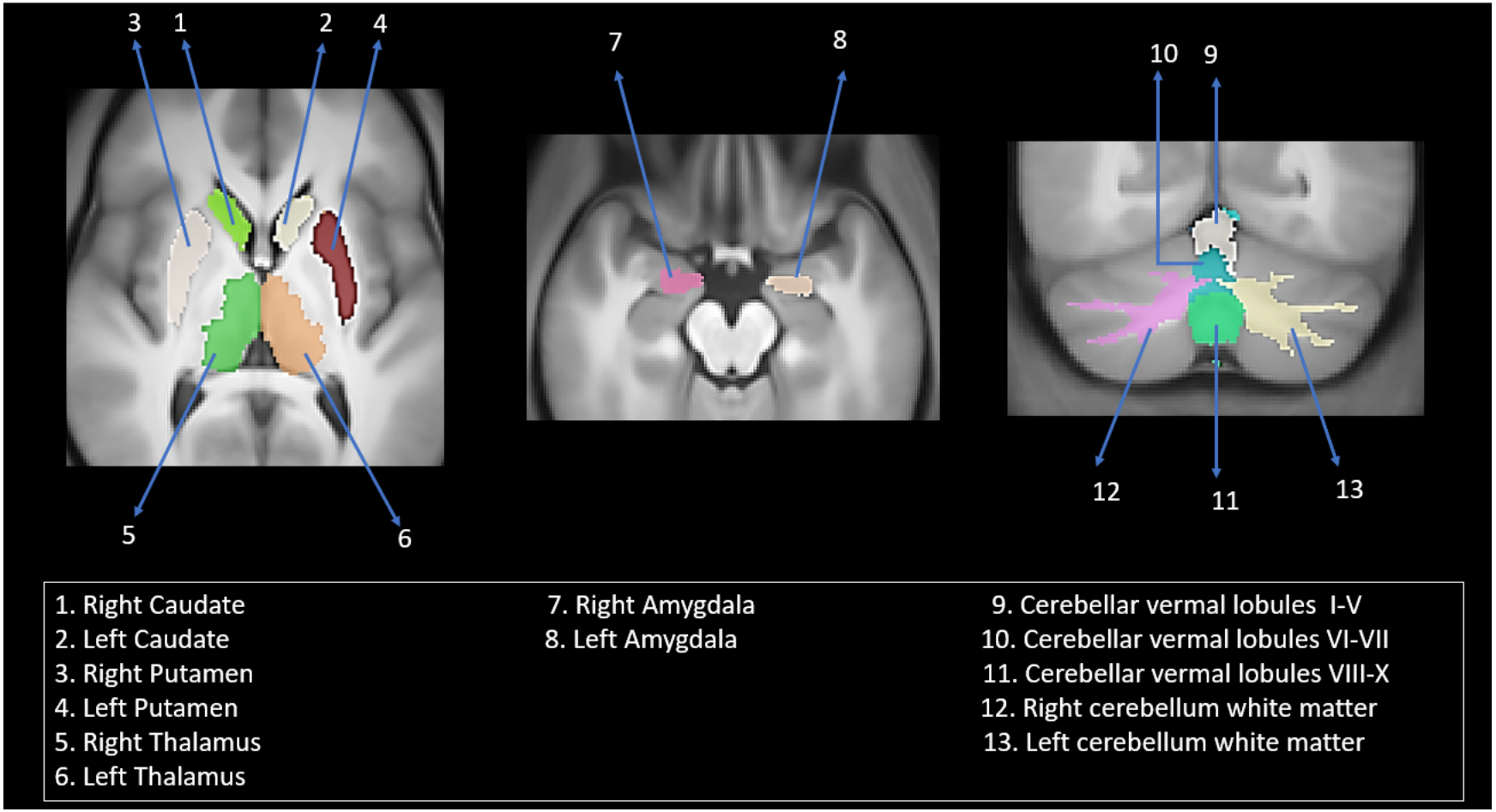


5. Splenium of corpus callosum (SCC)
17. Anterior Limb of Internal Capsule Right (ALIC)
18. Anterior Limb of Internal Capsule Left (ALIC)
19. Posterior Limb of Internal Capsule Right (PLIC_R)
20. Posterior Limb of Internal Capsule Left (PLIC_L)
23. Anterior Corona Radiata Right (ACR_R)
24. Anterior Corona Radiata Left (ACR_R)
27. Posterior Corona Radiata Right (PCR_R)
28. Posterior Corona Radiata Left (PCR_L)
35.Cingulum Cingulate Gyrus Right (CCG_R)
36. Cingulum Cingulate Gyrus Left (CCG_L)
47. Tapetum Right
48. Tapetum Left

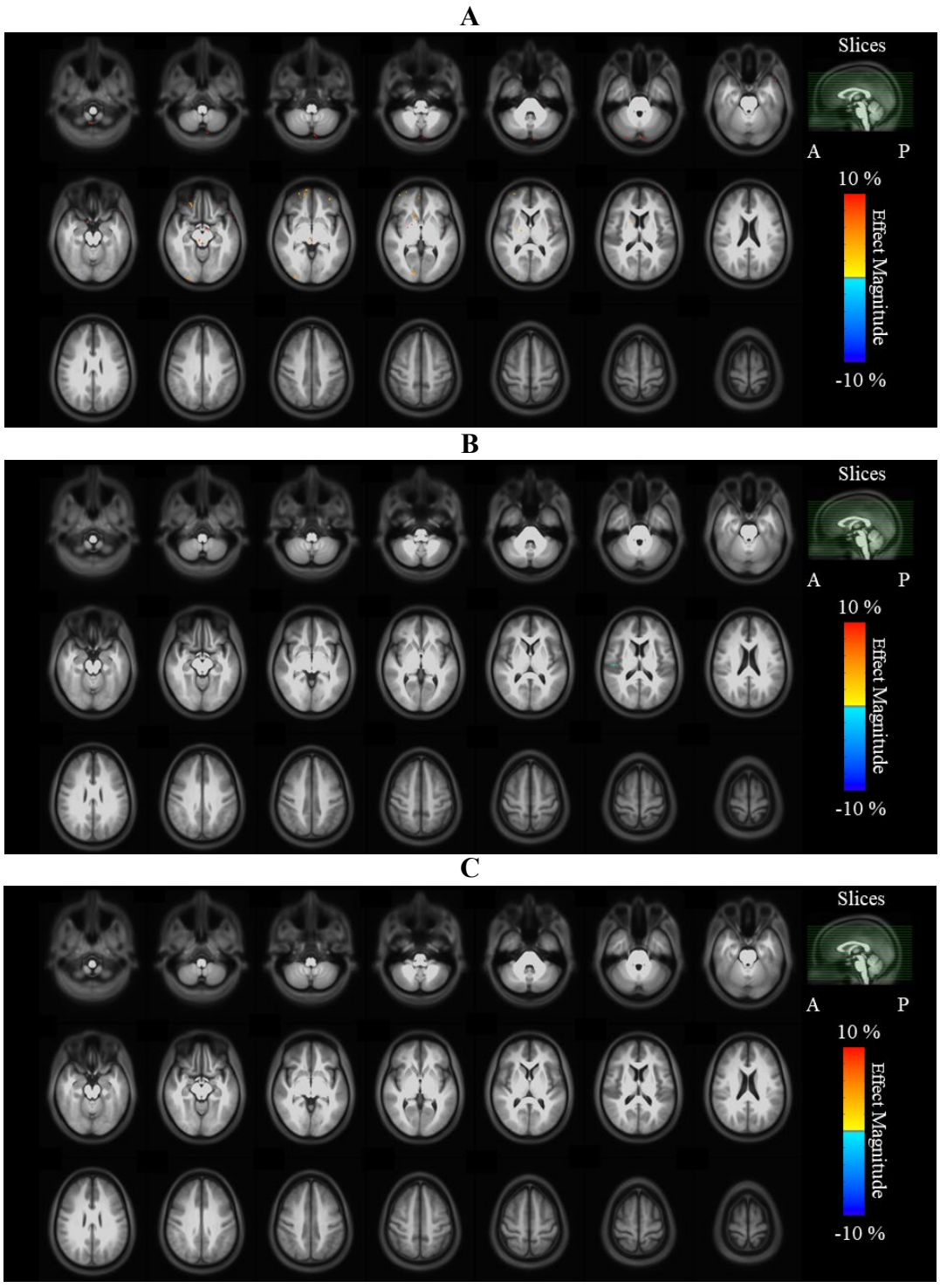




eFigure 1.**A-G** DTI white matter (WM) ROI location that are shown on the study template directionally encoded color (DEC) map. These ROIs are defined based on the JHU MNI white matter atlas. The ROIs were first defined on a connectome white matter template. The connectome template was registered to the study template using a tensor-based registration and the white matter ROIs were warped using the spatial transformations and nearest neighbor interpolation into the study template space.

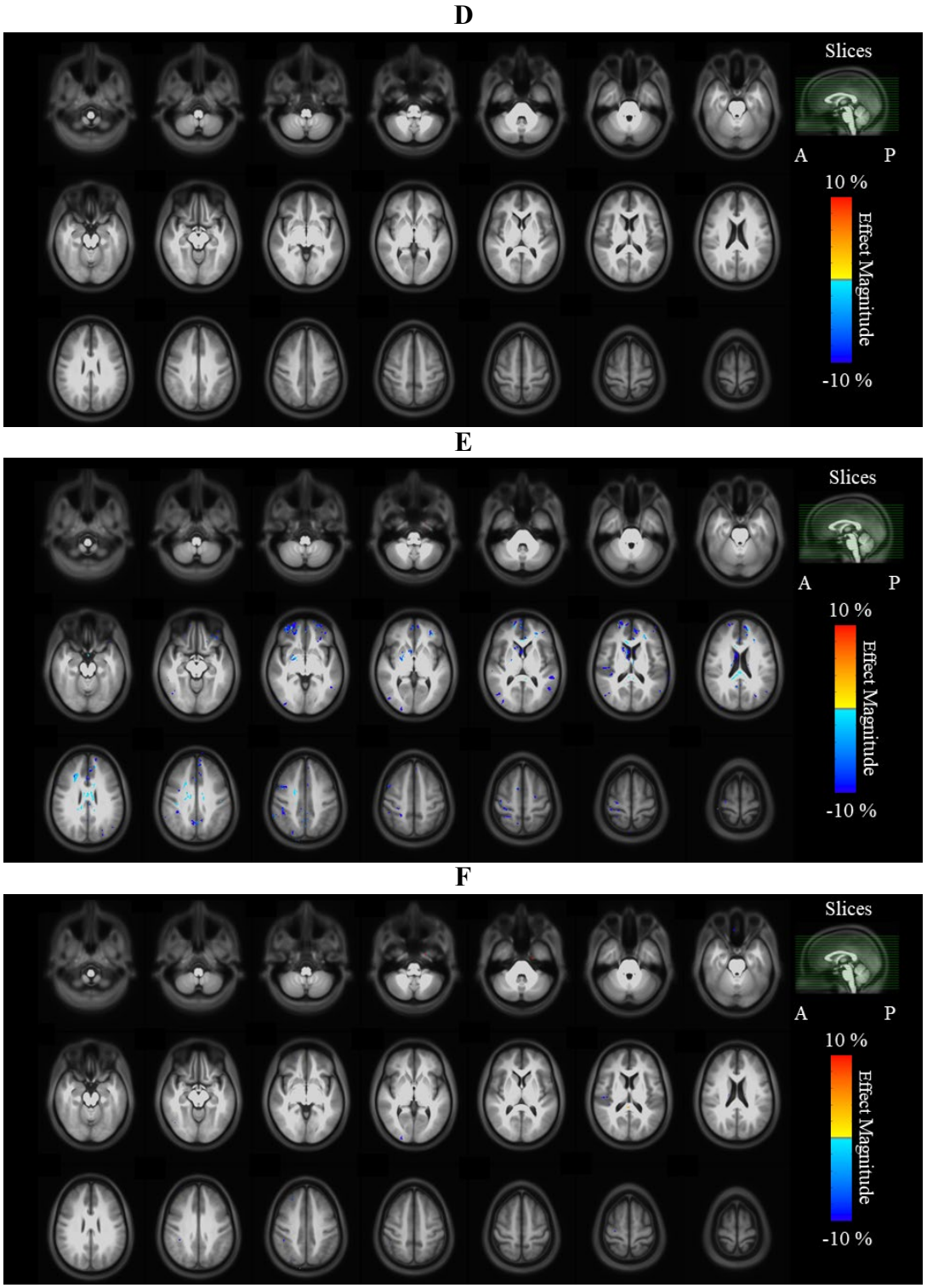


eFigure 2. Select GM and cerebellar WM ROIs used in the study.



eFigure 3.A Magnitude of difference maps from two-group whole brain voxel-wise comparisons between participants with AHI and controls for Par-MD. The magnitude of difference is only displayed for voxels that survive $P < 0.01$ for a two-sided test and not adjusted for multiple comparisons. The magnitude maps are superimposed on a T1W study template. The sagittal slice on the right panel with green lines traversing from anterior (A) to posterior (P), is a reference to the locations of the axial slices shown. Blue areas correspond to regions of significantly lower Par-MD in the participants with AHI than in the controls while red areas represent regions of higher Par-MD in the participants with AHI than in the controls. A few voxels show higher Par-MD in participants with AHI than controls, but they were small in both magnitude (<5%) and spatial extent. The description concerning the magnitude map display

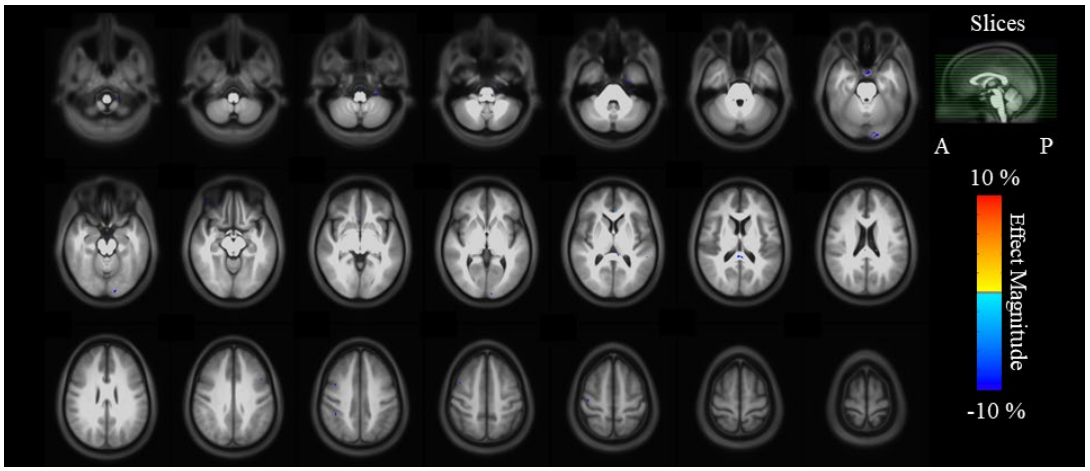
and image slices is applicable to the rest of the diffusion metrics shown in the subsequent figures. **B** Magnitude of difference maps from two-group whole brain voxel-wise comparisons for the CSF-SF. Differences between participants with AHI and controls for CSF-SF were small in both magnitude and of no apparent anatomical pattern. **C** No regions with significantly altered AD were observed in the brain parenchyma.



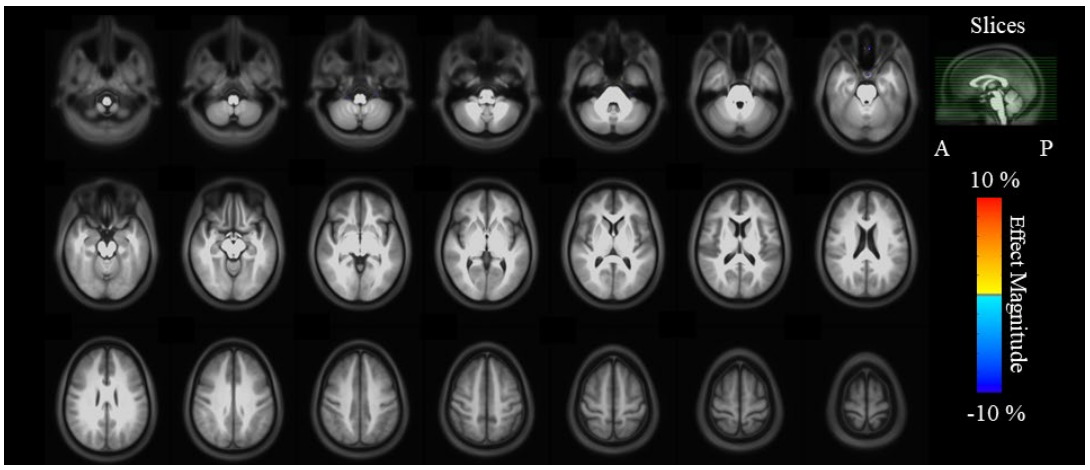
eFigure 3.D Magnitude of difference maps from two-group whole brain voxel-wise comparisons for RD. No regions with significantly altered RD were observed in the brain parenchyma. **E** Two-group whole brain voxel-wise comparisons reveal lower PA in

participants with AHI compared to controls in the corpus callosum (CC) (~ 2-3% in magnitude), and in regions located at the interfaces between gray matter (GM) and sulci (>=10% in magnitude), but this could arise from inconsistent inter-subject registration. **F** Few voxels in Splenium of CC showed slightly higher NG in participants with AHI compared to controls.

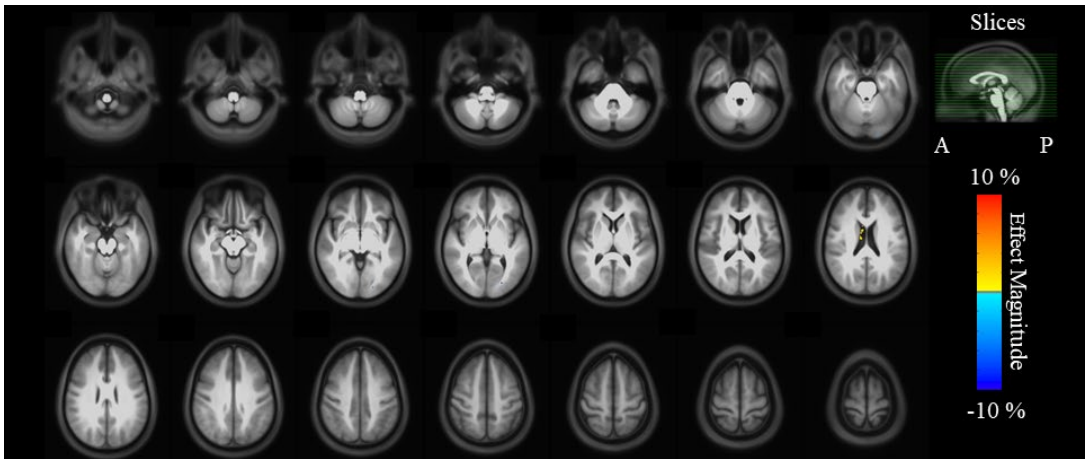
G



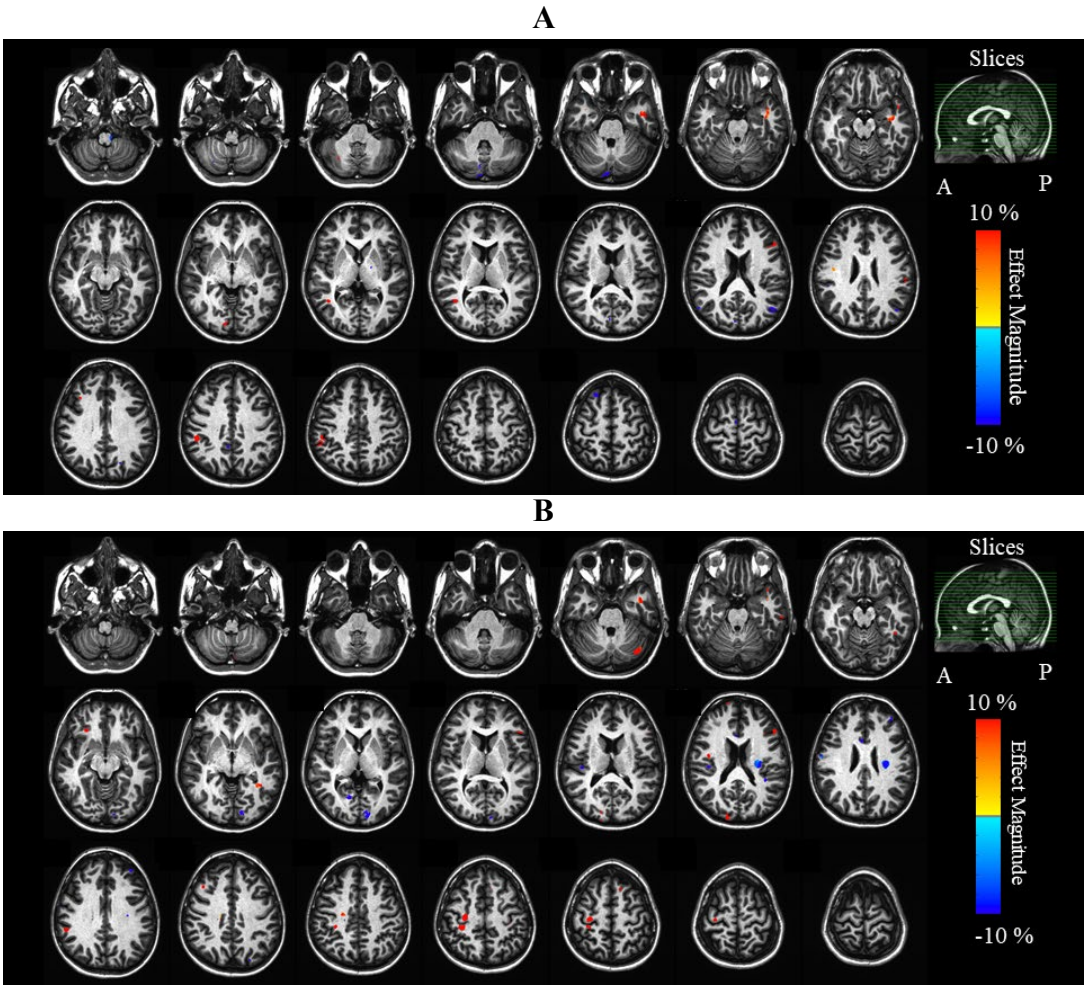
H



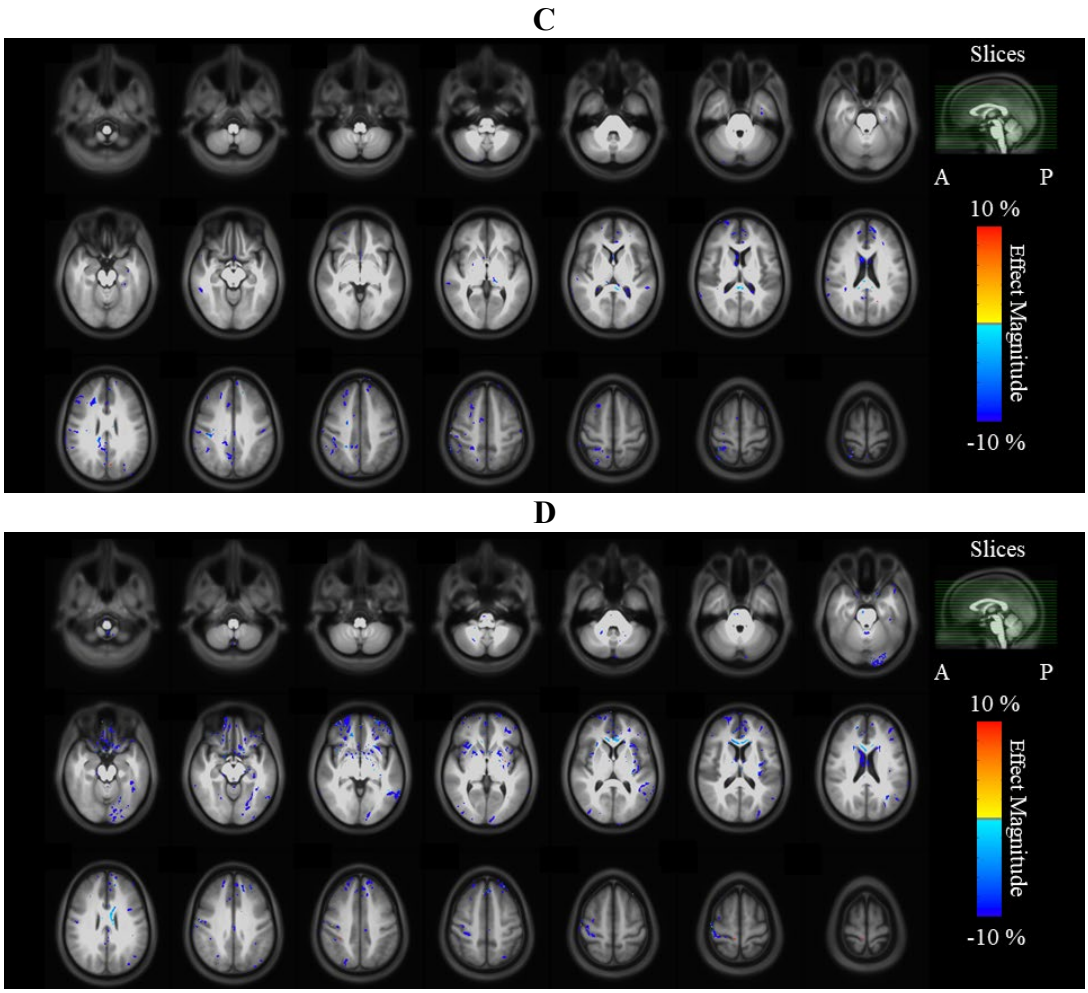
I



eFigure 3.G Magnitude of difference maps from two-group whole brain voxel-wise comparisons for RTAP. A few voxels showed lower RTAP in splenium of CC in participants with AHI than in controls. **H** and **I** However, no regions with significantly altered RTOP and RTPP, respectively, were observed in the brain parenchyma.

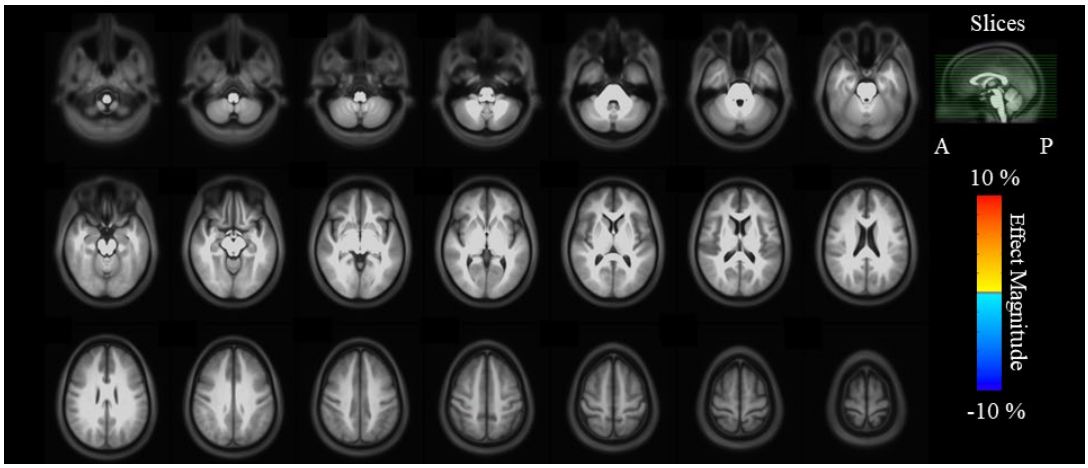


eFigure 4.A Magnitude of difference maps from the three-group voxel-wise comparisons for volumetrics between AHI1 vs controls, and **B** AHI2 vs controls. All voxels displayed, survived $P < 0.01$ for a two-sided test, not adjusted for multiple comparisons. Blue areas correspond to regions of significantly smaller volume in the participants with AHI than in the controls while red areas represent regions of larger volume in the participants with AHI than in the controls. A few regions with significantly higher and lower volumes (magnitude of difference ~8-10% or higher in either directions) were observed in the AHI1 and AHI2 subgroups compared to controls, with no clear anatomical pattern or bilateral consistency.

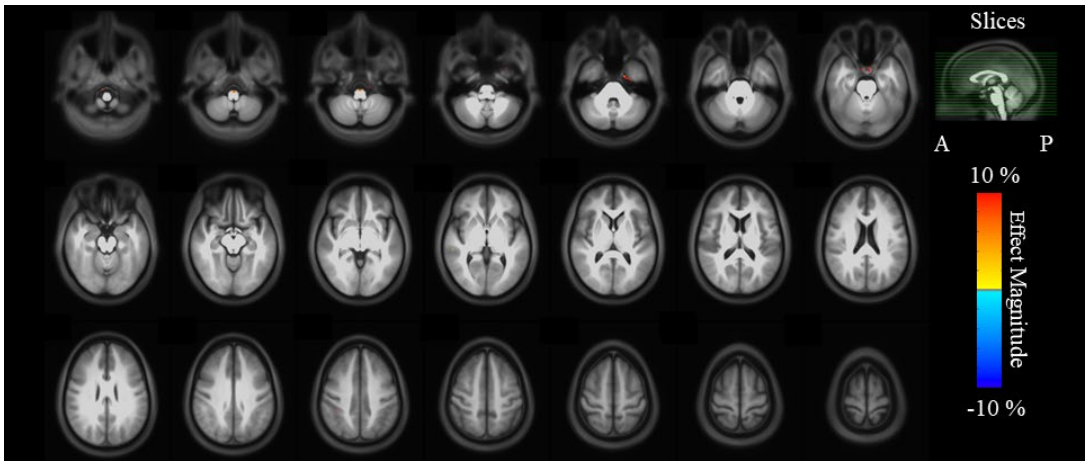


eFigure 4.C Magnitude of difference maps from the three-group whole brain voxel-wise comparisons for FA between AHI1 vs controls, and **D** AHI2 vs controls. All displayed voxels survived $P < 0.01$ for a two-sided test, from an unadjusted TFCE statistical map. AHI1 and AHI2 participants showed significantly lower FA than controls in the corpus callosum (CC) (magnitude ~2-3%), and in some regions located at the interfaces between gray matter (GM) and sulci which could arise from inconsistent inter-subject registration. No regions with significantly higher FA in participants with AHI were observed. The following figures for the three-group comparison of diffusion metrics will follow the same theme.

E

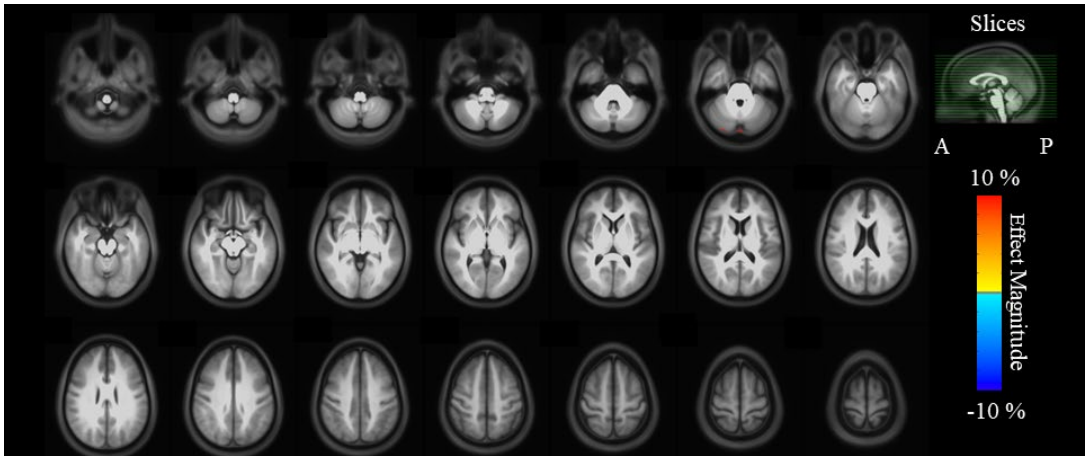


F

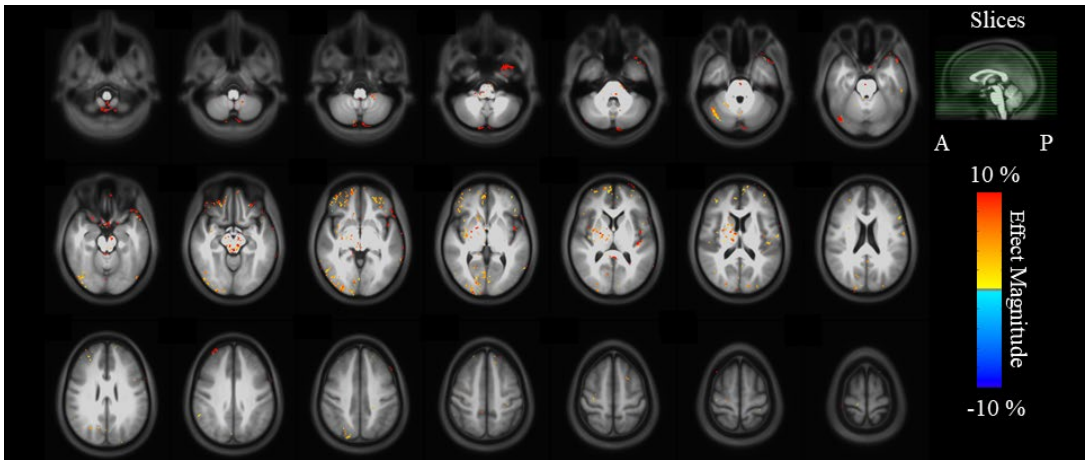


eFigure 4.E Magnitude of difference maps from three-group whole brain voxel-wise comparisons for MD between AHI1 vs controls, and **F** AHI2 vs controls. No regions with significantly altered MD were observed in the brain parenchyma for participants within AHI1 and AHI2 categories compared to controls.

G

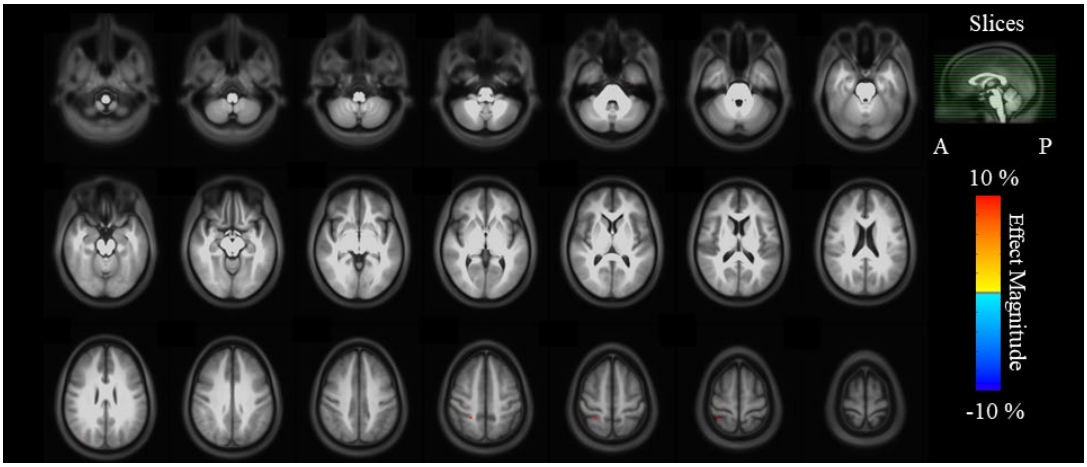


H

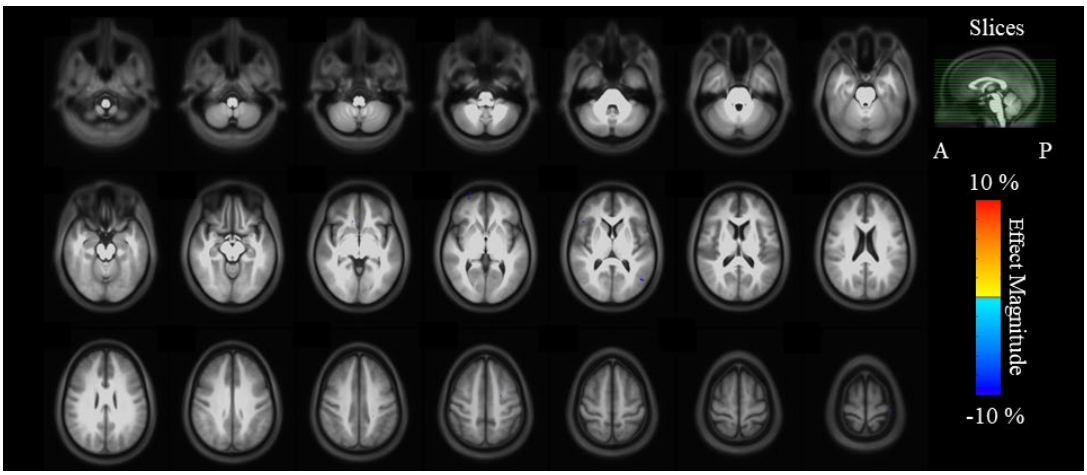


eFigure 4.G Magnitude of difference maps from three-group whole brain voxel-wise comparisons for Par-MD between AHI1 vs controls, and **H** AHI2 vs controls. There were a few regions with significantly higher Par-MD (magnitude ≥ 2-10%) in the participants with AHI2 category compared to controls but with no clear anatomical patterns.

I

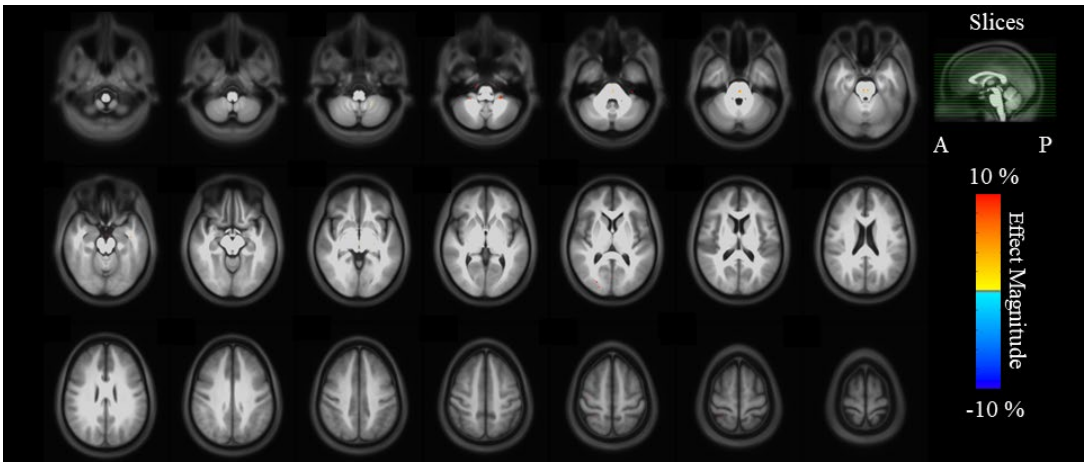


J

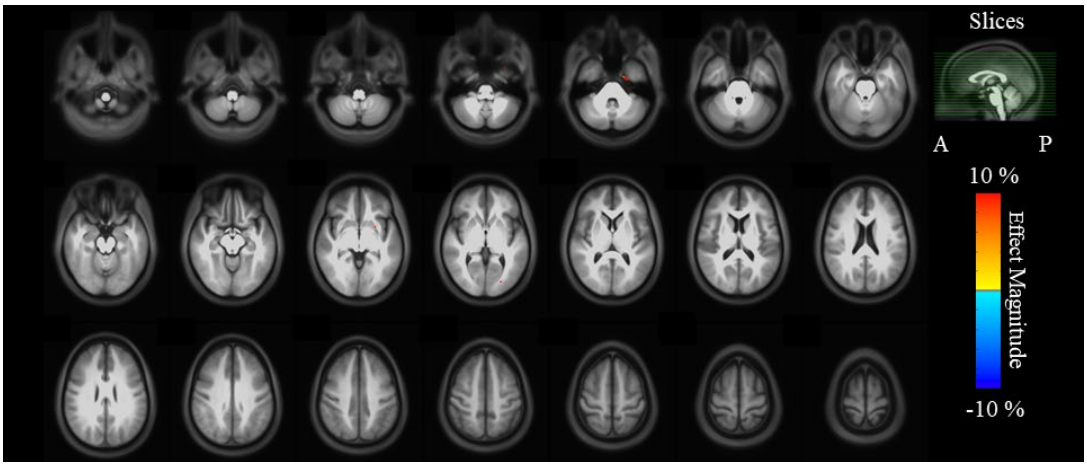


eFigure 4.I Magnitude of difference maps from the three-group whole brain voxel-wise comparisons for CSF-SF between AHI1 vs controls, and **J** AHI2 vs controls. No regions with significantly altered CSF-SF were observed in the brain parenchyma for participants within AHI1 and AHI2 categories compared to controls.

K

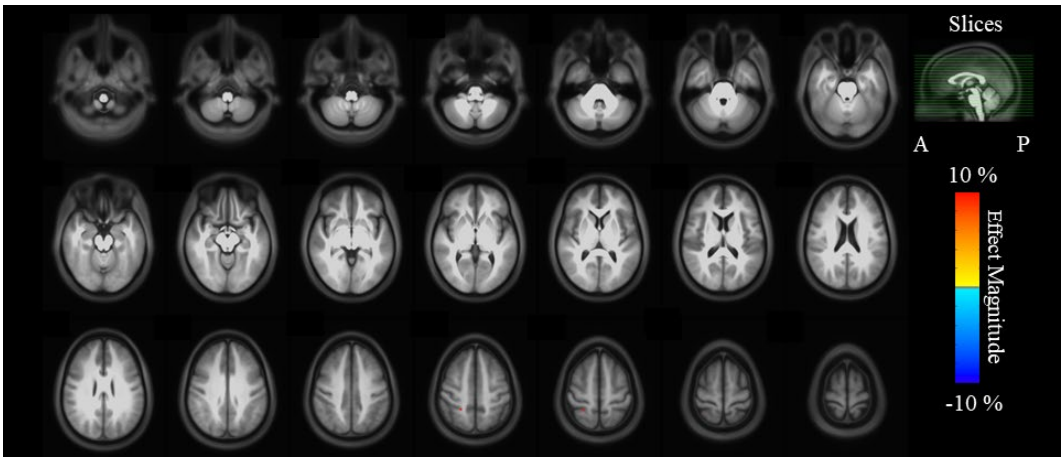


L

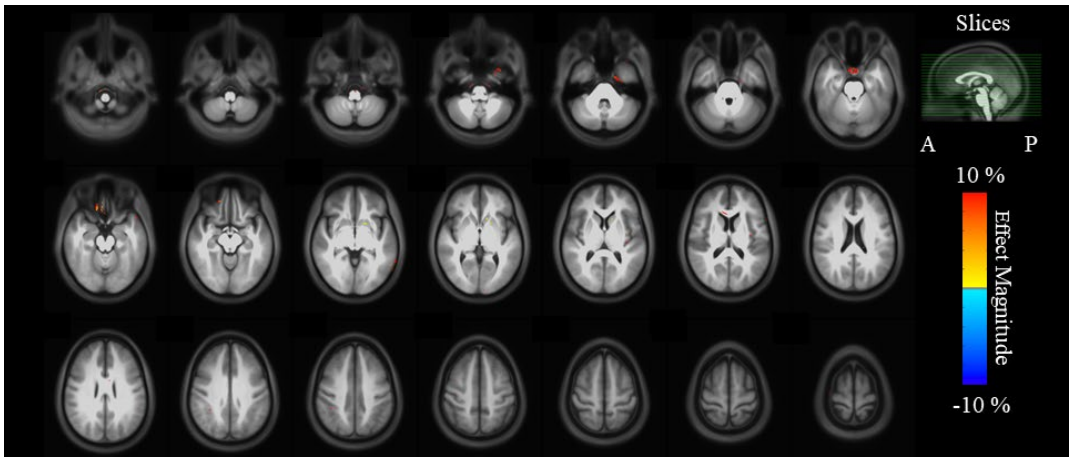


eFigure 4.K Magnitude of difference maps from three-group whole brain voxel-wise comparisons for AD between AHI1 vs controls, and **L** AHI2 vs controls. A few regions were observed with significantly higher AD (magnitude ≤10%) for participants within AHI1 category compared to controls. However, they were small in both magnitude and spatial extent with no apparent anatomical pattern.

M

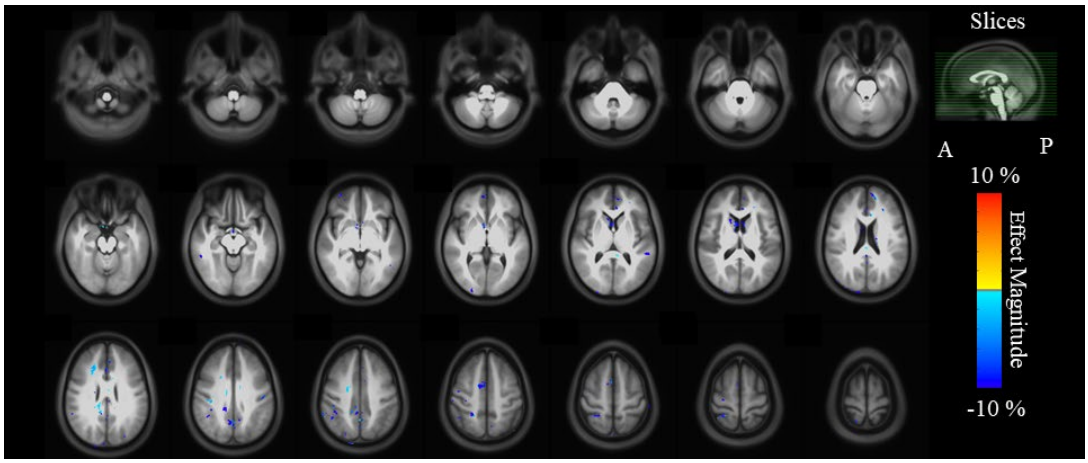


N

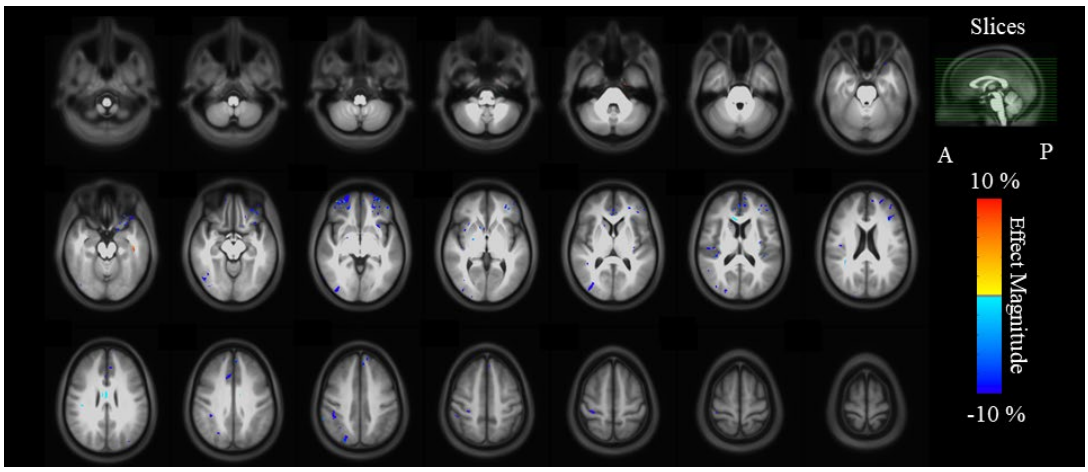


eFigure 4.M Magnitude of difference maps from the three-group whole brain voxel-wise comparisons for RD between AHI1 vs controls, and **N** AHI2 vs controls. When compared to controls, AHI2 showed higher RD in a few voxels in Genu of the CC, however, no other regions exhibited significantly altered RD in the brain parenchyma or had a left -right consistency for any of the subgroups.

O

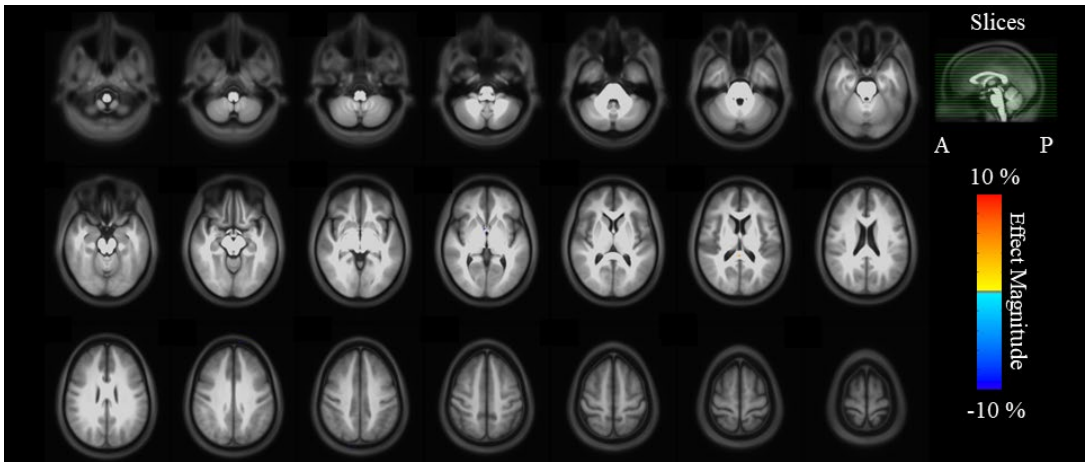


P

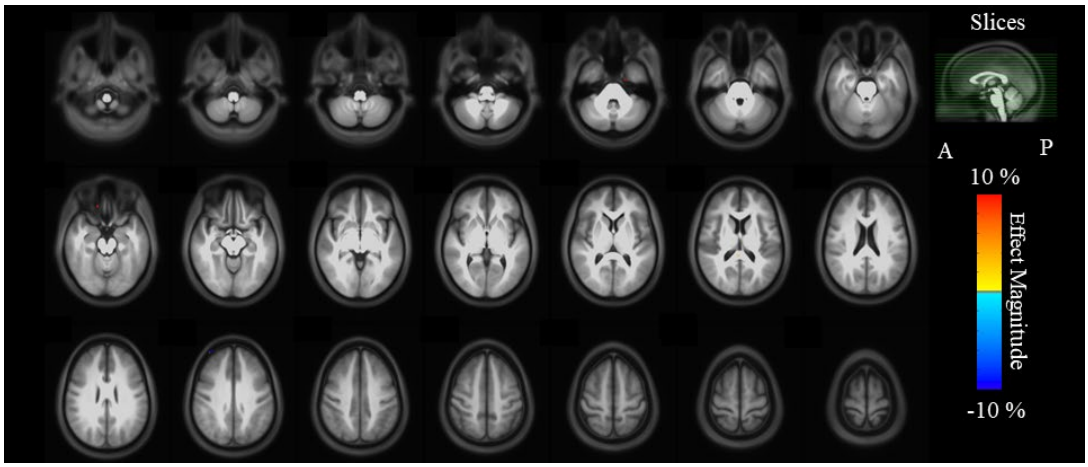


eFigure 4.O Magnitude of difference maps from the three-group whole brain voxel-wise comparisons for PA between AHI1 vs controls and **P** AHI2 vs controls. Participants within AHI 1 and AHI2 categories exhibited significantly lower PA than controls in the corpus callosum (CC), ~ 2-3% in magnitude, and in regions located at the interfaces between gray matter (GM) and sulci, ~8% in magnitude, which could arise from inconsistent inter-subject registration. No regions with significantly higher PA in participants with AHI were observed.

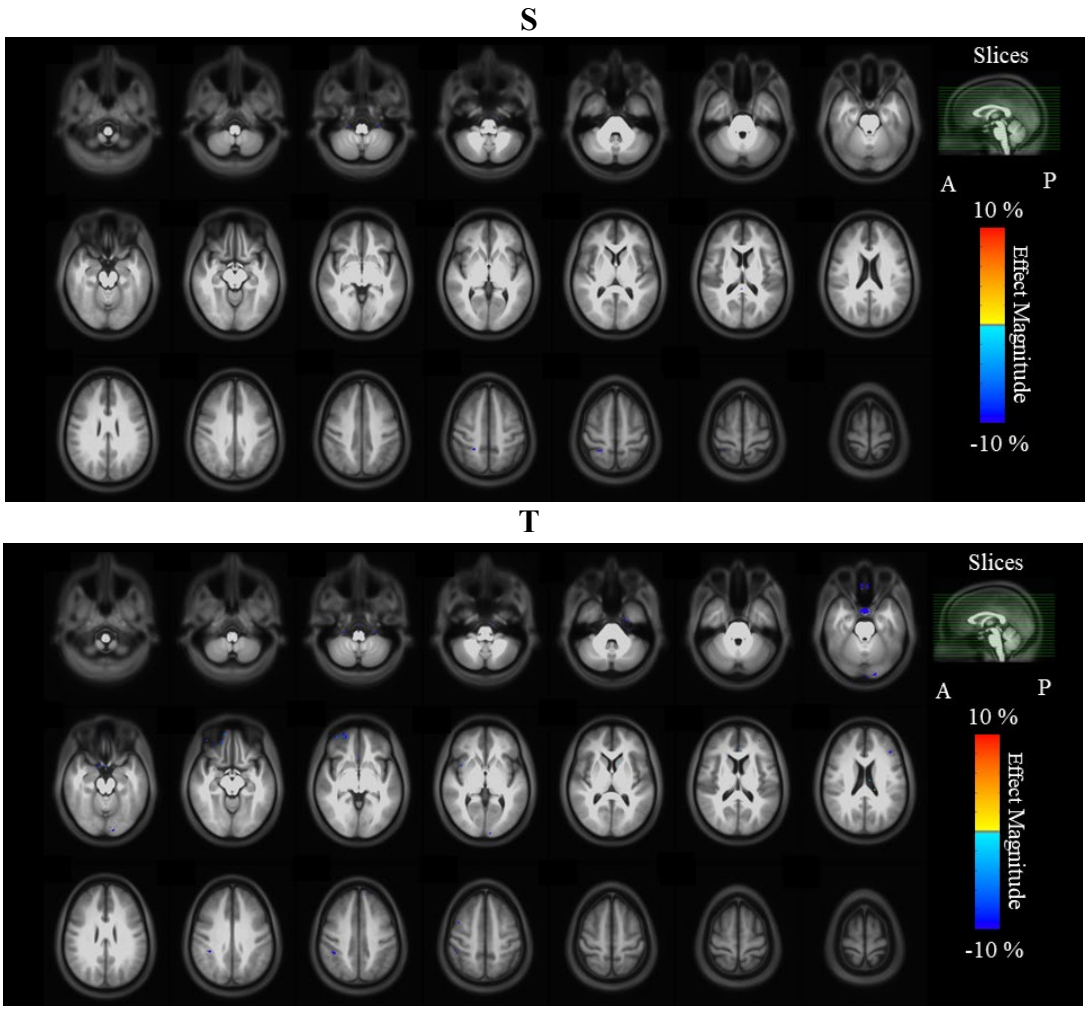
Q



R

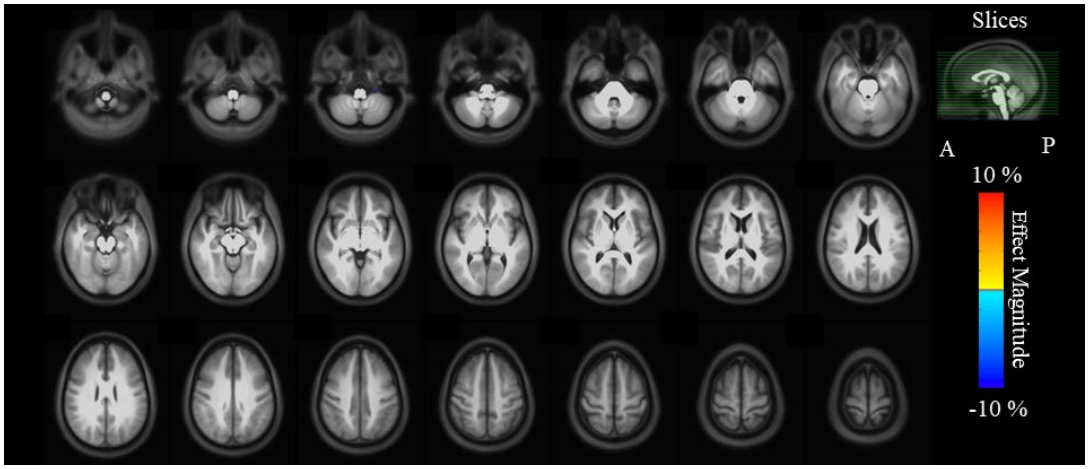


eFigure 4.Q Magnitude of difference maps from the three-group whole brain voxel-wise comparisons for NG between AHI1 vs controls, **R** AHI2 vs controls. No regions with significantly altered NG were observed in the brain parenchyma for participants within AHI1 and AHI2 categories compared to controls.

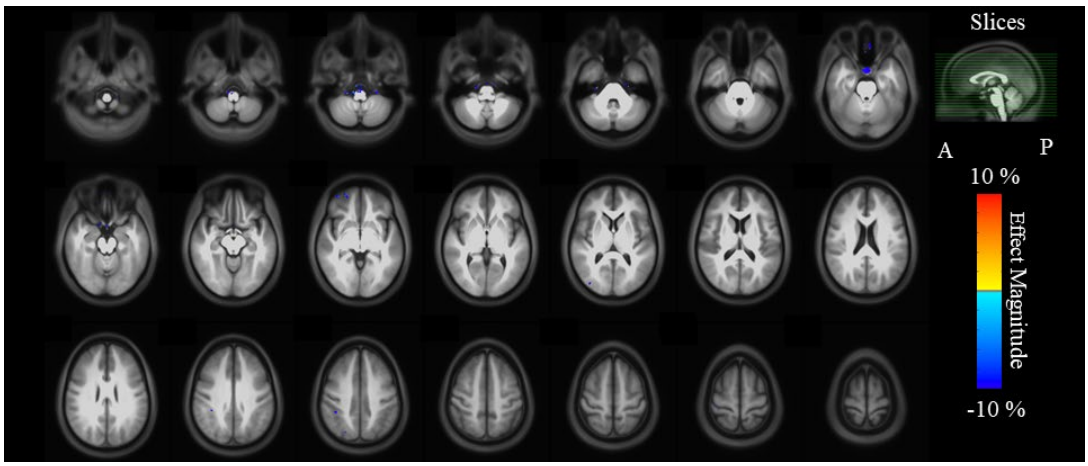


eFigure 4.S Magnitude of difference maps from the three-group whole brain voxel-wise comparisons for RTAP between AHI1 vs controls, and **T** AHI2 vs controls. The participants within the AHI1 category showed lower RTAP in the Splenium of the CC than controls. A few regions showed lower RTAP than controls in participants within the AHI2 category, however with no apparent anatomical pattern or bilateral consistency.

U

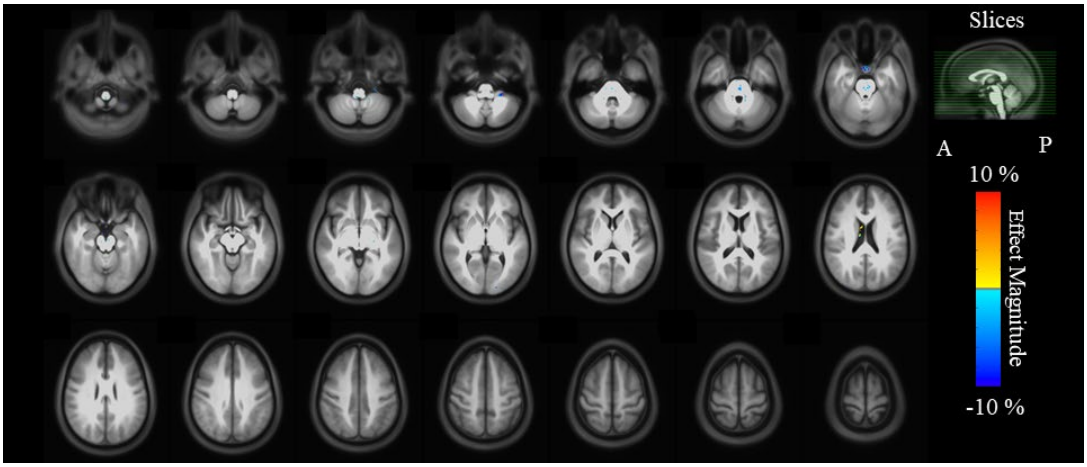


V

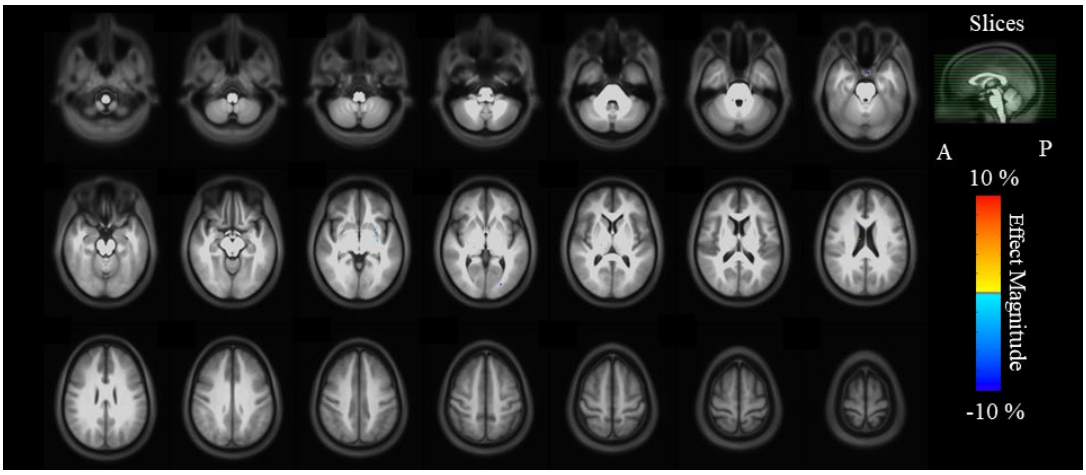


eFigure 4.U Magnitude of difference maps from the three-group whole brain voxel-wise comparisons for RTOP between AHI1 vs controls, and **V** AHI2 vs controls. Some regions with significantly lower RTOP can be observed in participants within AHI2 category compared to controls, but with no clear anatomical pattern or left-right consistency.

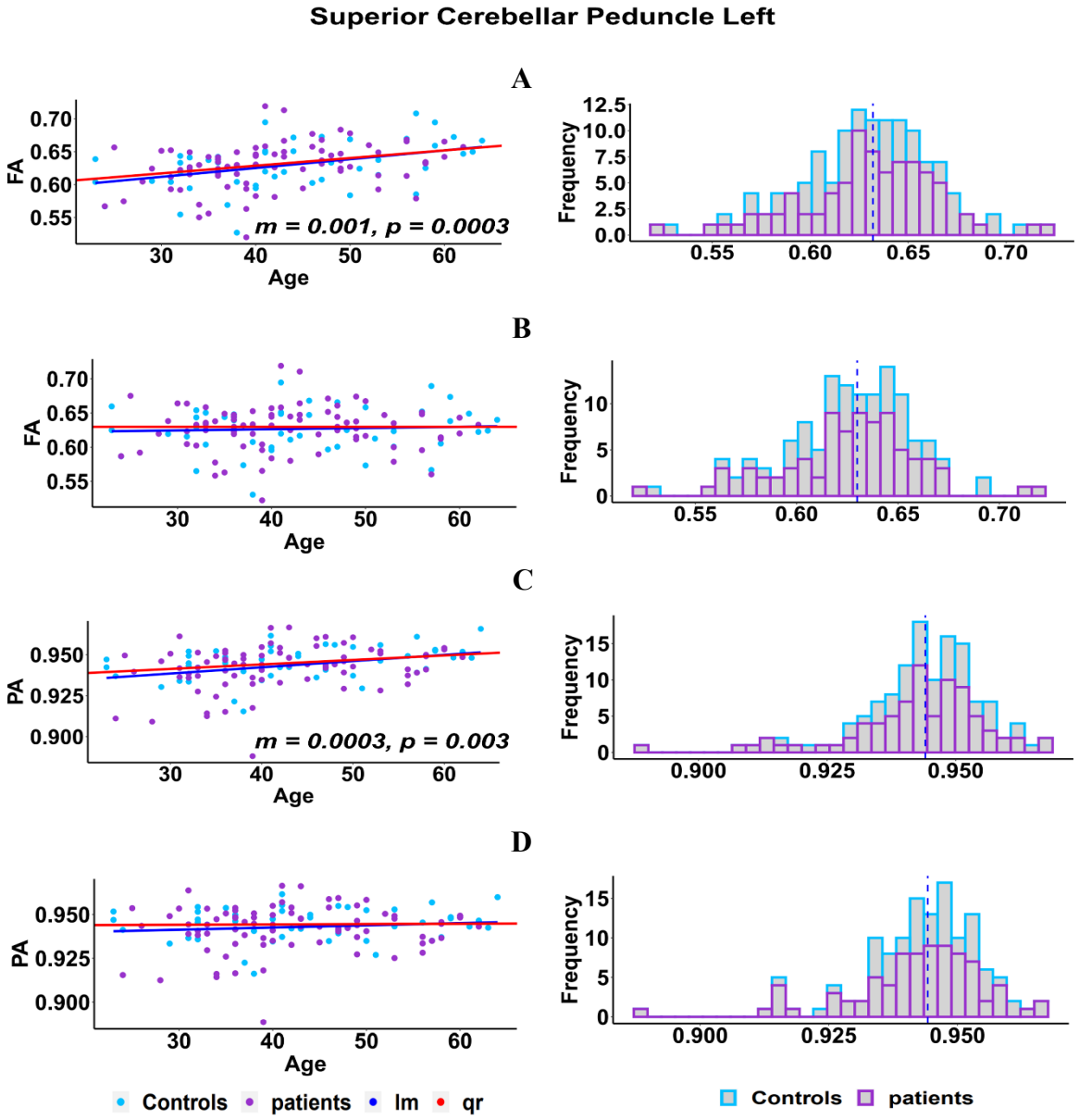
W



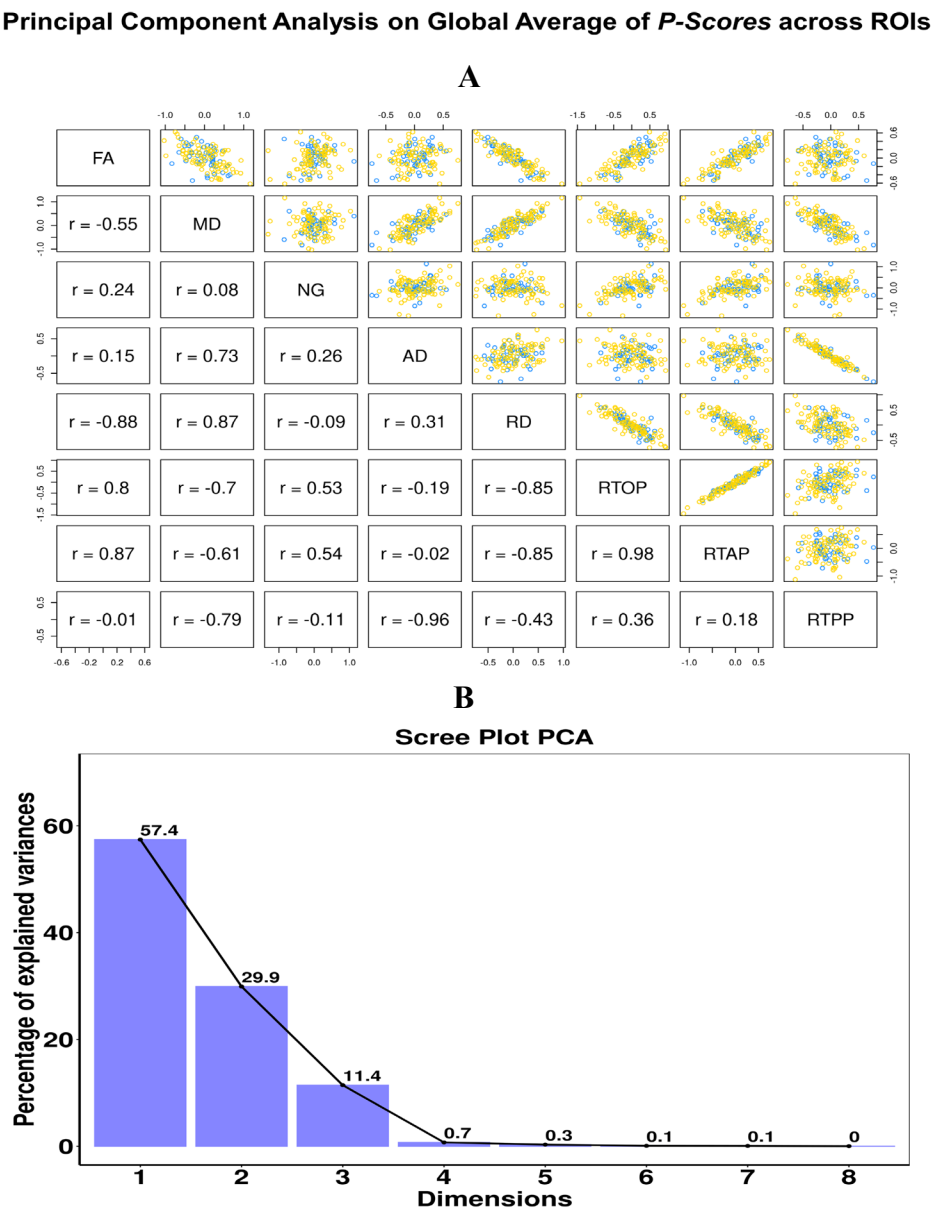
X



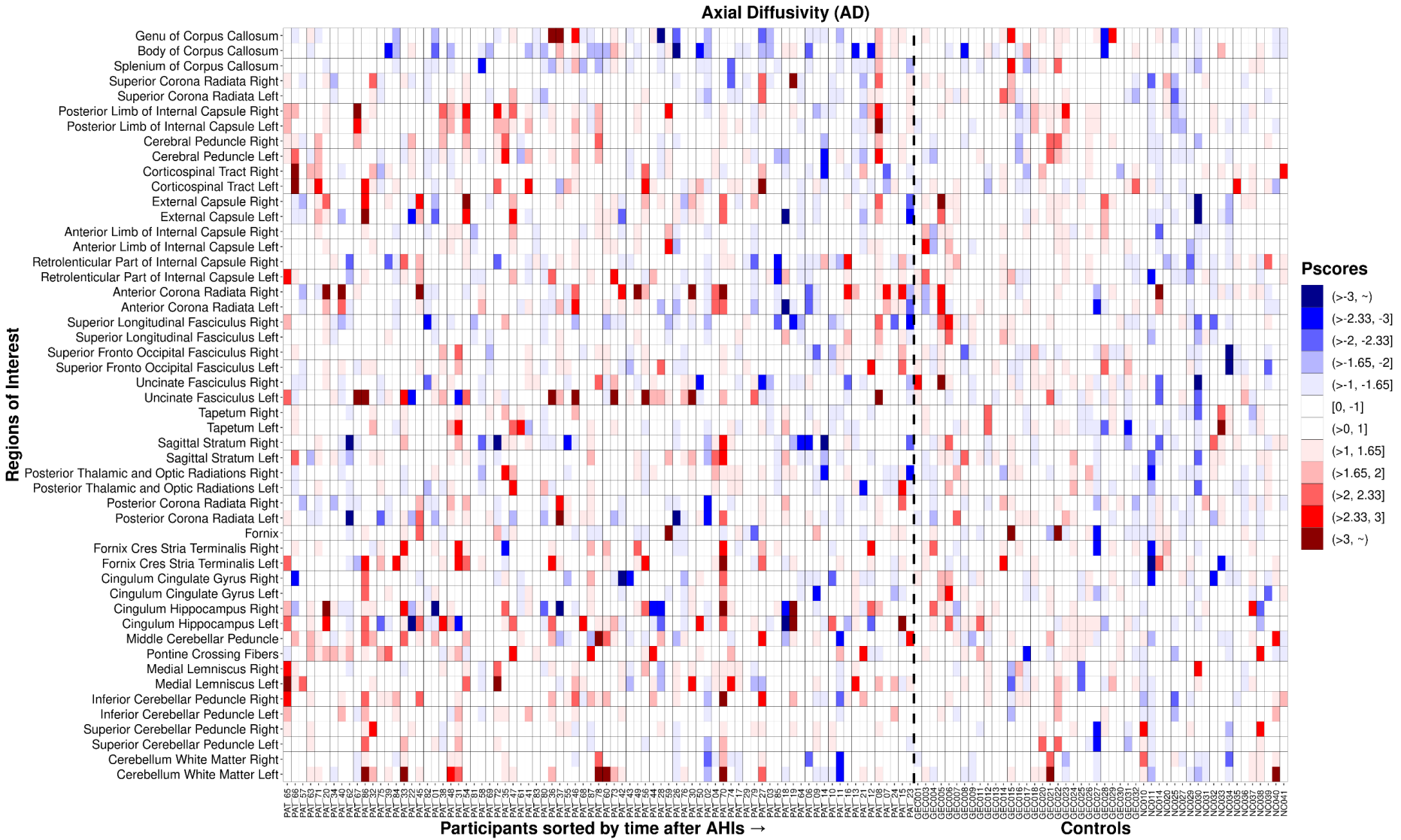
eFigure 4.W Magnitude of difference maps from the three-group whole brain voxel-wise comparisons for RTPP between AHI1 vs controls, and **X** AHI2 vs controls. A few voxels showed lower RTPP among participants in AHI1 and AHI2 categories compared to controls. However, these differences were small in both magnitude and spatial extent with no apparent anatomical pattern.



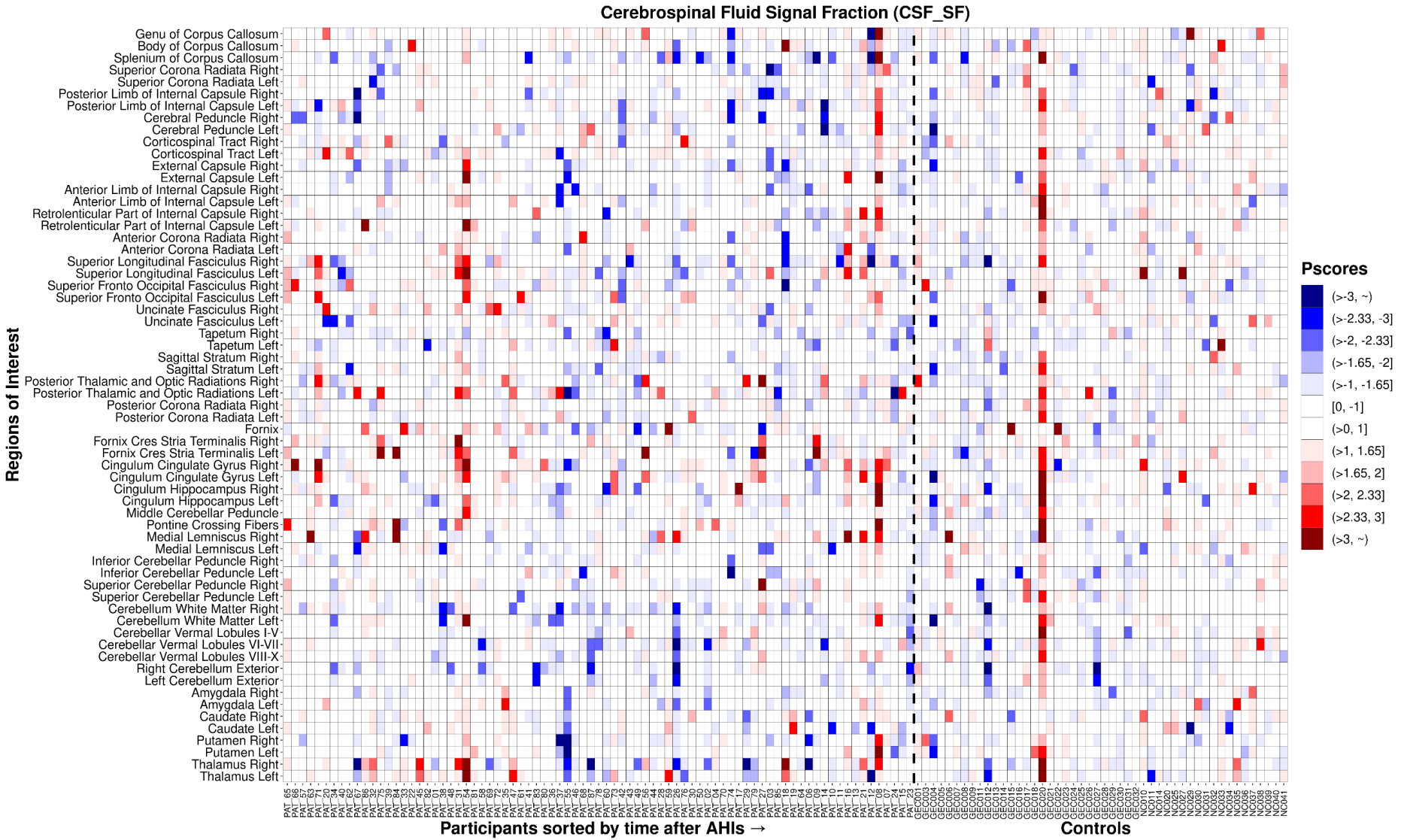
eFigure 5. Demonstrating median age-fitting using quantile regression. **A** *Left:* Significant effect of age within the ROI: Superior Cerebellar Peduncle Left for the diffusion metric - FA. The *"m"* represents the slope estimate of age, and the *"p"* represents the corresponding P value for significance obtained from the regression model. The solid red line is the quantile regression fit for the median age. The solid blue line represents the linear fit, representing the conventional mean regression model for age. *Right:* The distributions of the two sample groups appear to be slightly heavy tailed. The dotted line represents the overall median across the two groups. **B** *Left*: No linear trend across the sample groups after removing the age effect and adding the intercept, modelled to represent FA at the median sample age of *41 years. Right:* The corresponding distribution after removing age effects and a minor shift of the median line to the left can be observed. **C-D** Similar to **A** and **B** but for the DTI metric – PA. Within the same ROI, the distribution is heavily skewed for PA and quantile regression of the median age provides the more accurate model for age-fitting compared to the mean regression model which assumes a normal distribution.



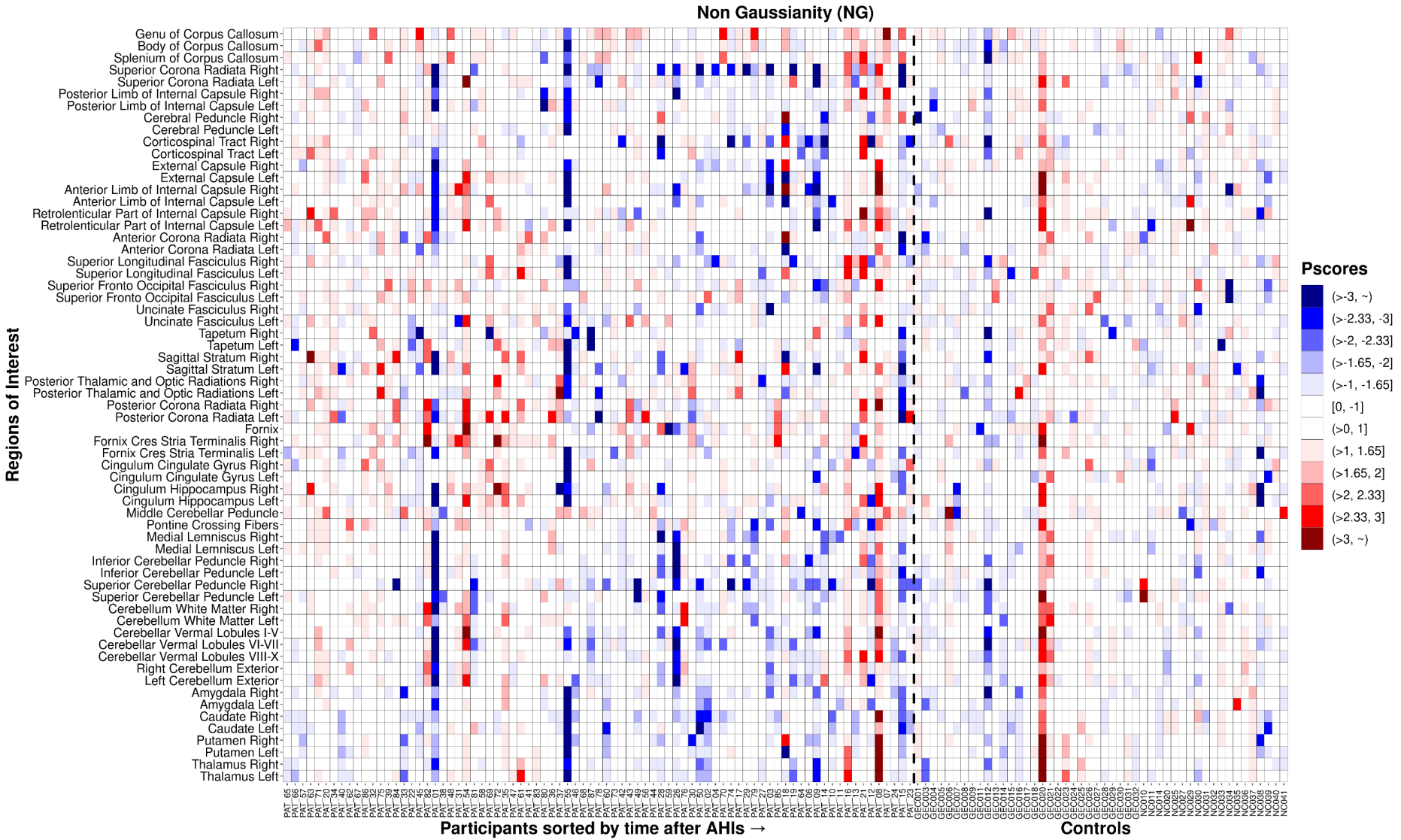
eFigure 6. Demonstrating collinearity between variables and percentage of variance explained by the principal components (PCs). **A** The correlation matrix with the scatter plot of observations in the upper panel and corresponding Pearson's *r* values on the lower panel. The blue circles represent controls, and the yellow circles represent the participants with AHIs. One can easily observe that there is a range of weak to moderate to very strong collinearity between these variables, e.g., FA has a strong negative association with RD but very weak with RTPP, whereas a strong positive relationship with RTOP and RTAP. PCA is very suitable under these circumstances as it will generate components that will be orthogonal to each other, thereby, effectively accounting for the collinearities. **B** A scree plot showing the percentage of variance explained by the eight PCs. The first two PCs explain > 87% of the variance, therefore, these two PCs are suitable candidates in leveraging retention of the maximum information to summarize the data.



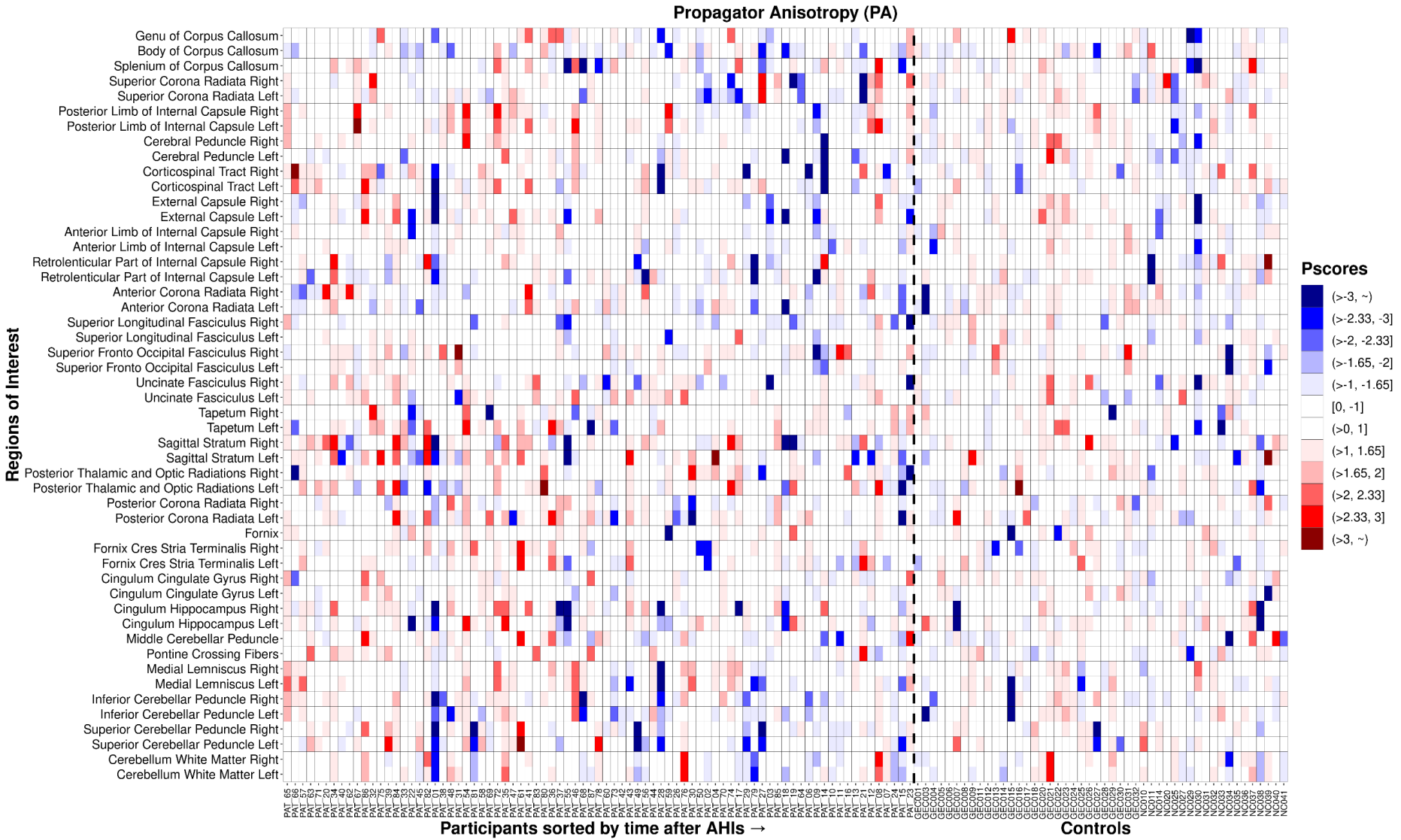
eFigure 7.A Heatmap of AD for participants with AHI and controls. The format is same as Figure 3 in the main manuscript and consistent for the rest of the figures in this set.



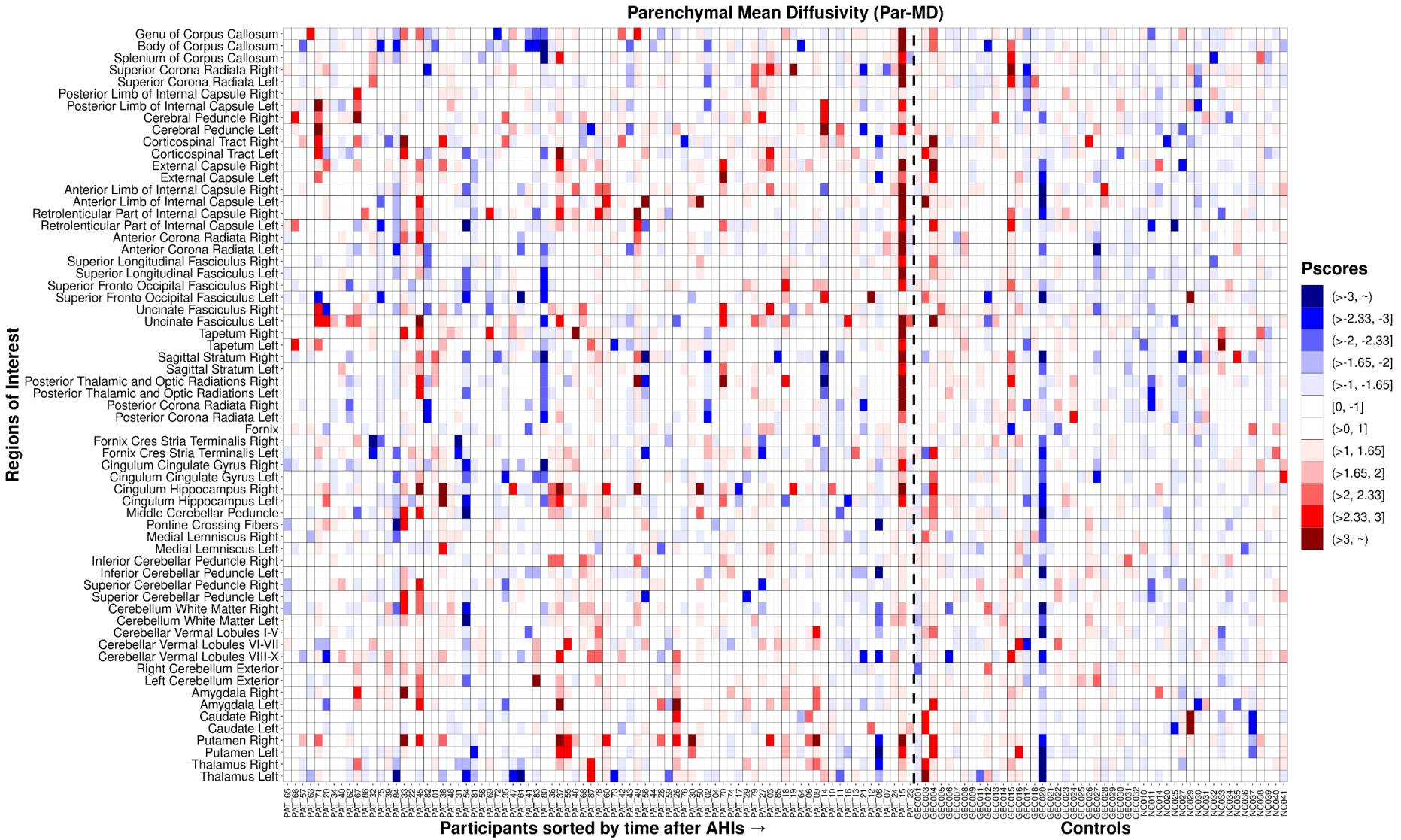
eFigure 7.B Heatmap of CSF-SF for participants with AHI and controls.



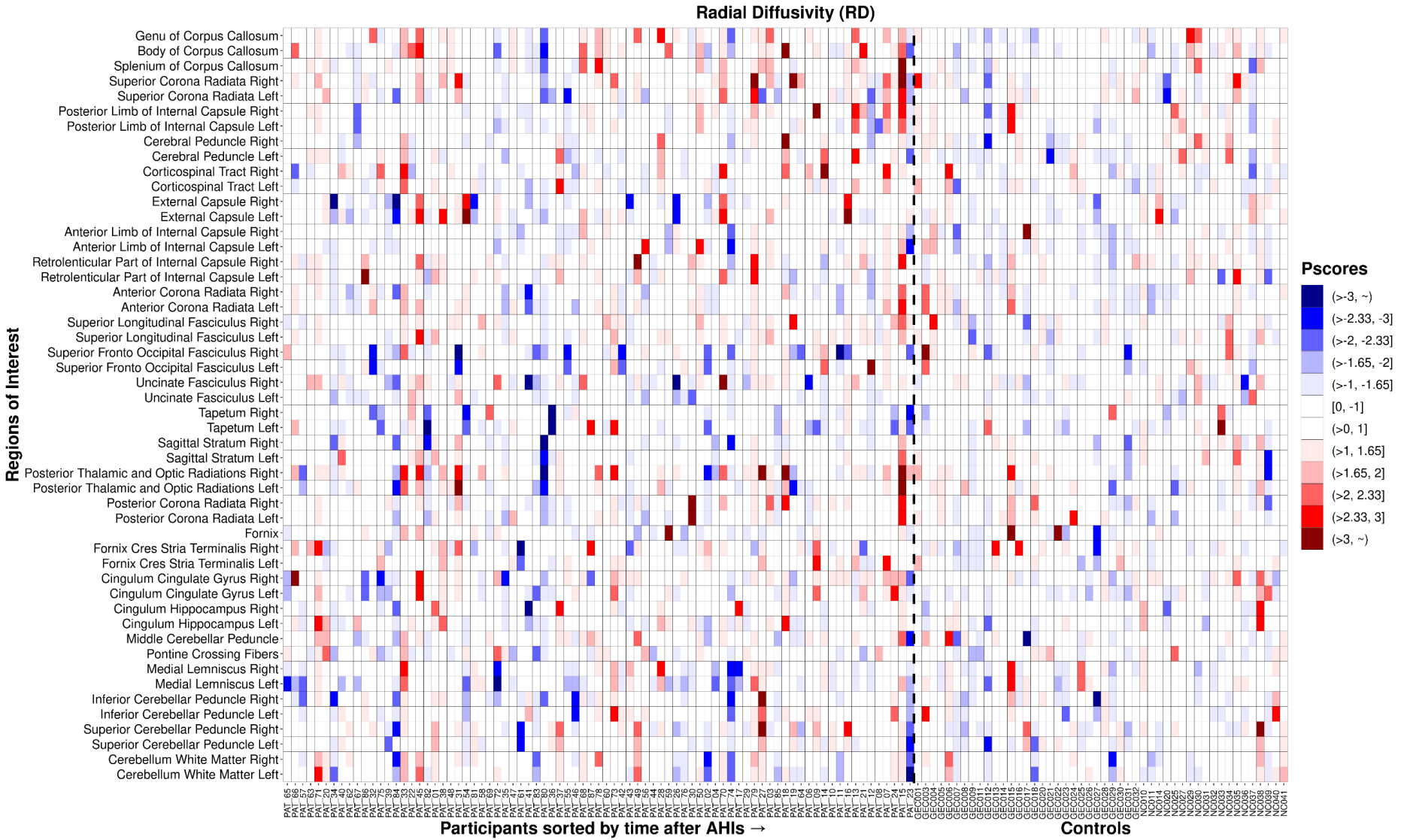
eFigure 7.C Heatmap of NG for participants with AHI and controls.



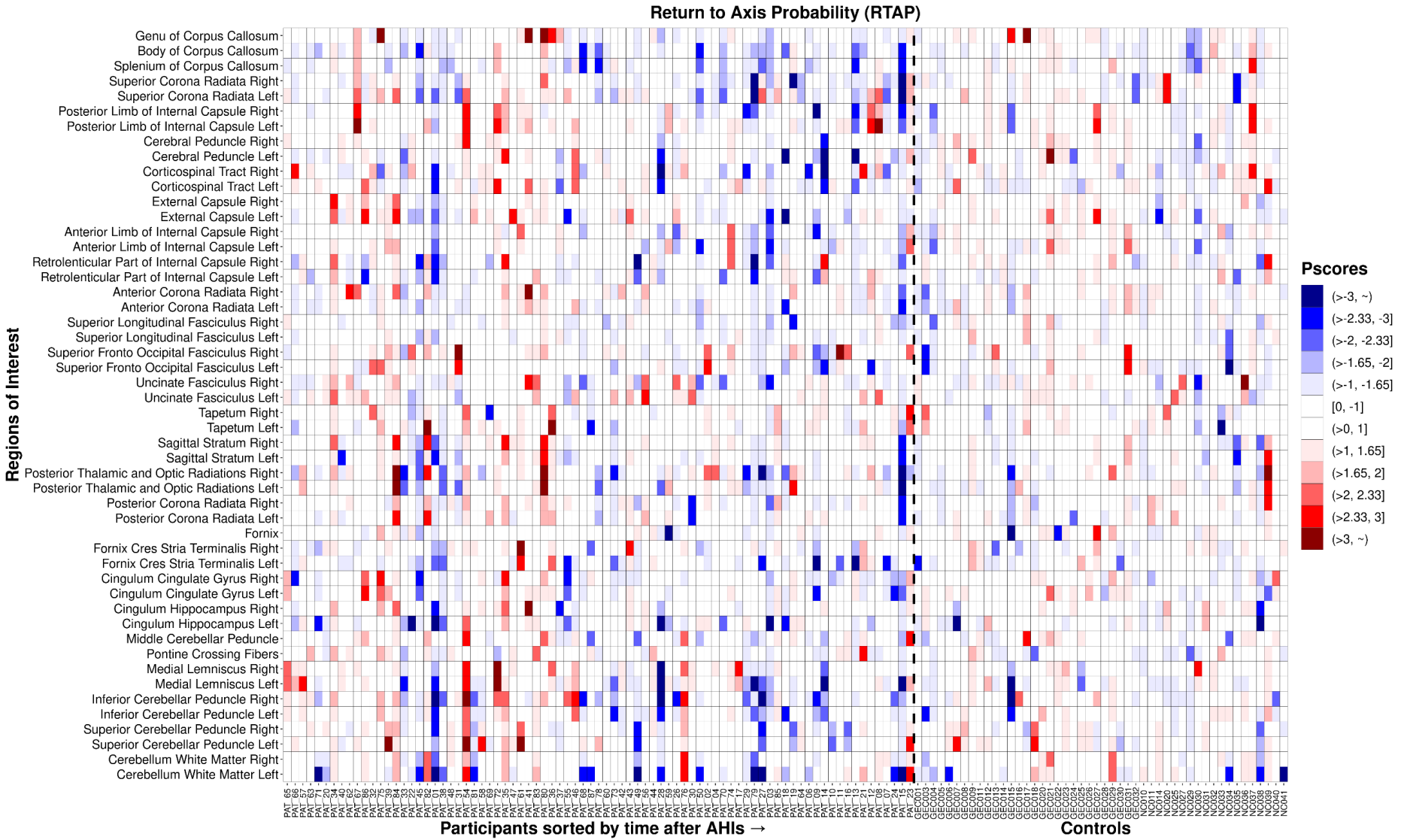
eFigure 7.D Heatmap of PA for participants with AHI and controls.



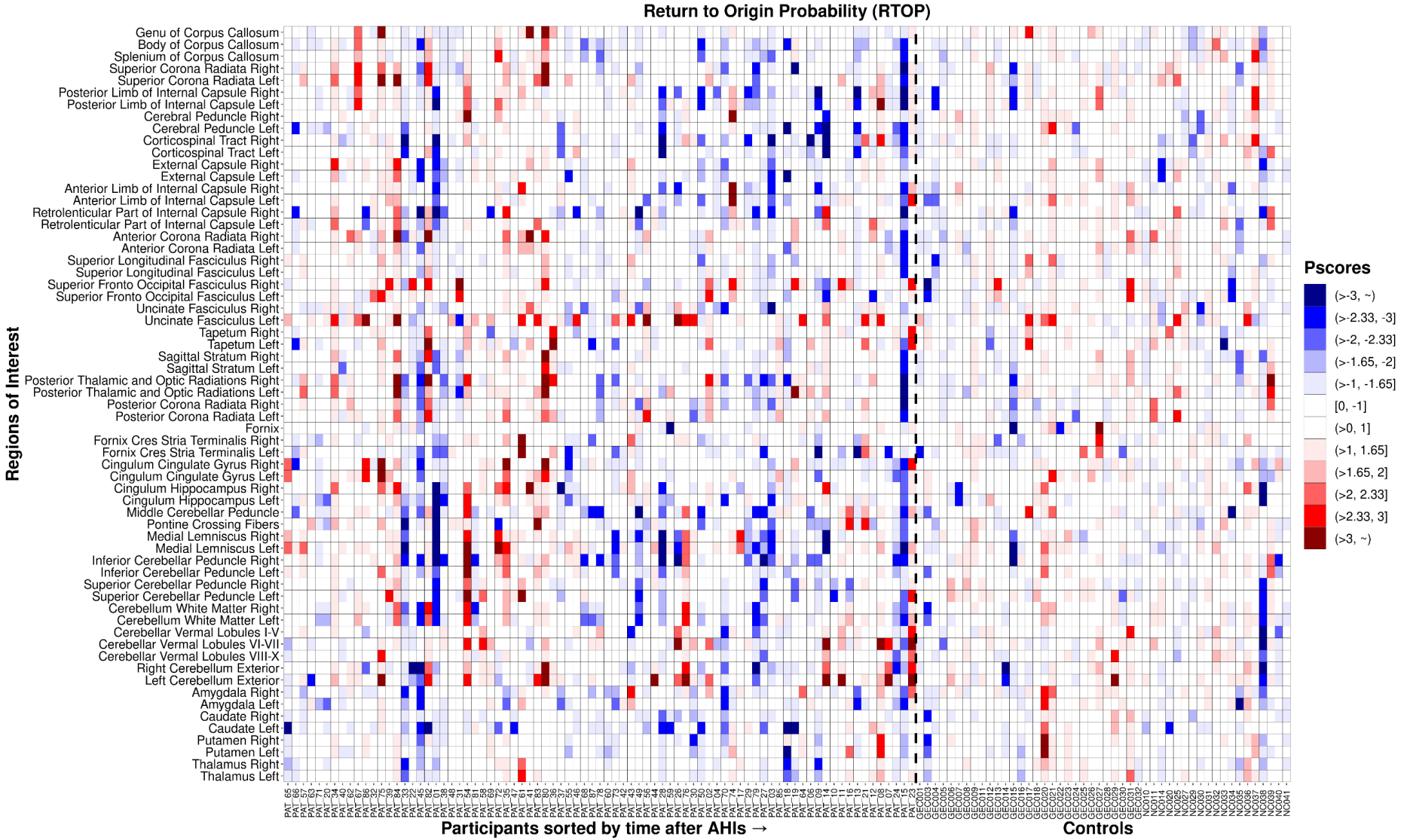
eFigure 7.E Heatmap of Par-MD for participants with AHI and controls.



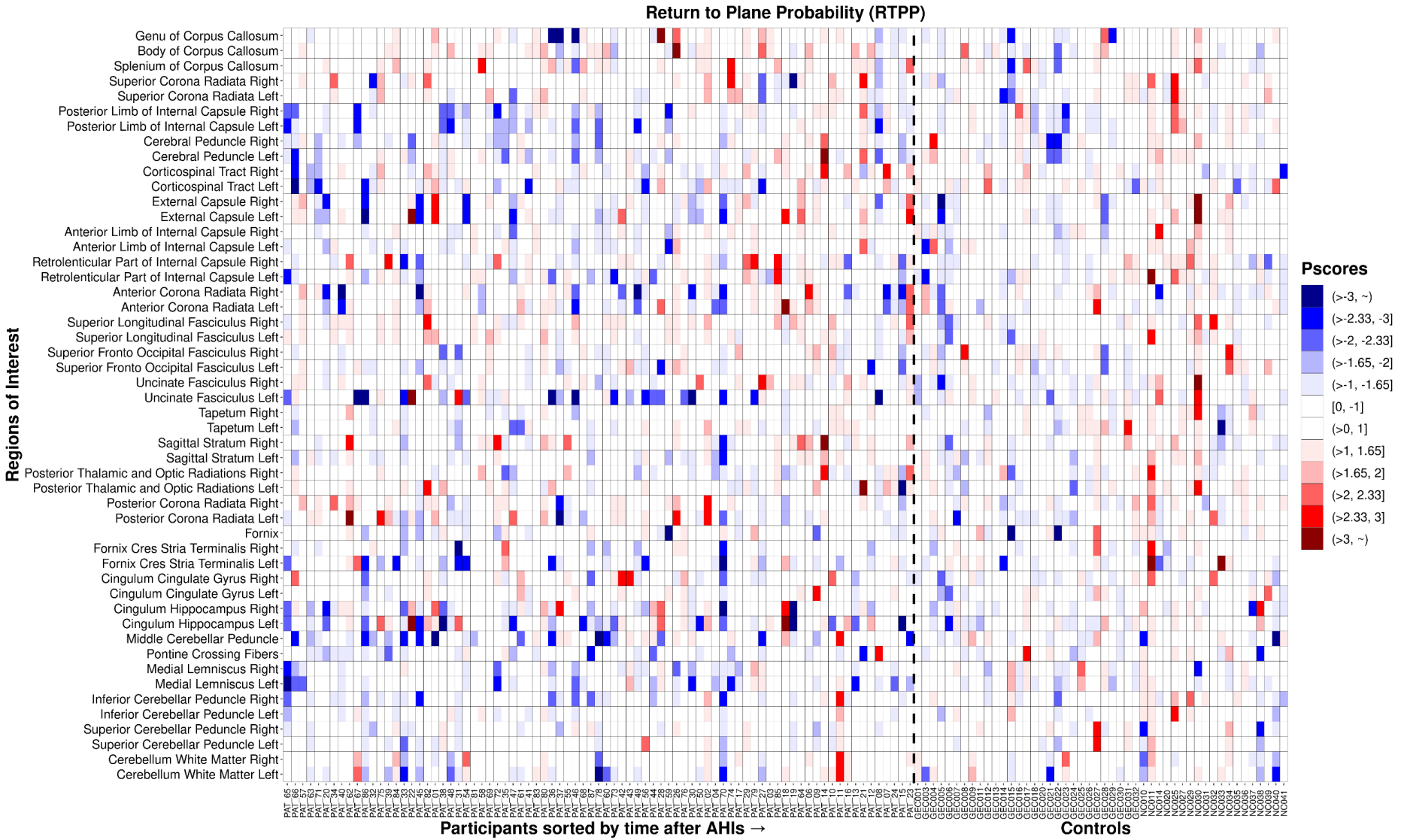
eFigure 7.F Heatmap of RD for participants with AHI and controls.



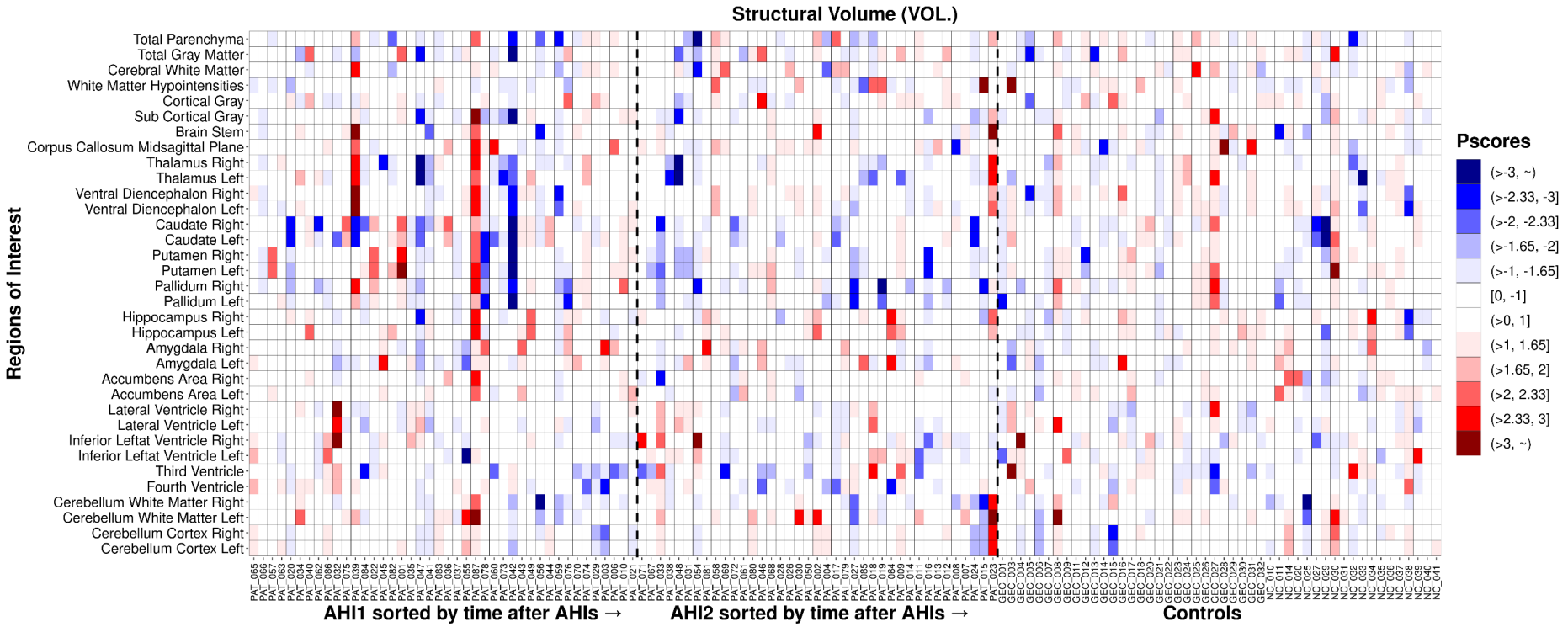
eFigure 7.G Heatmap of RTAP for participants with AHI and controls.



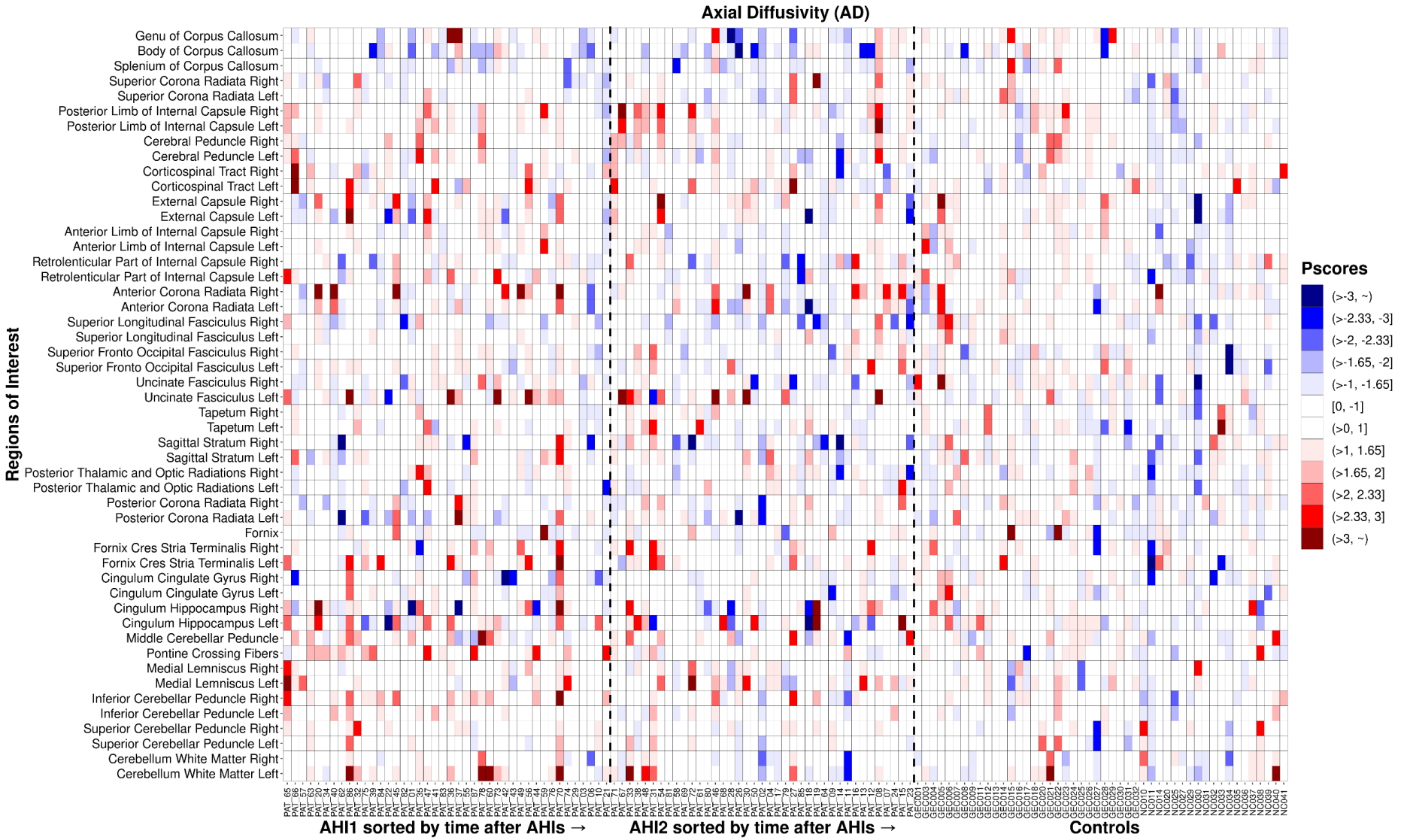
eFigure 7.H Heatmap of RTOP for participants with AHI and controls.



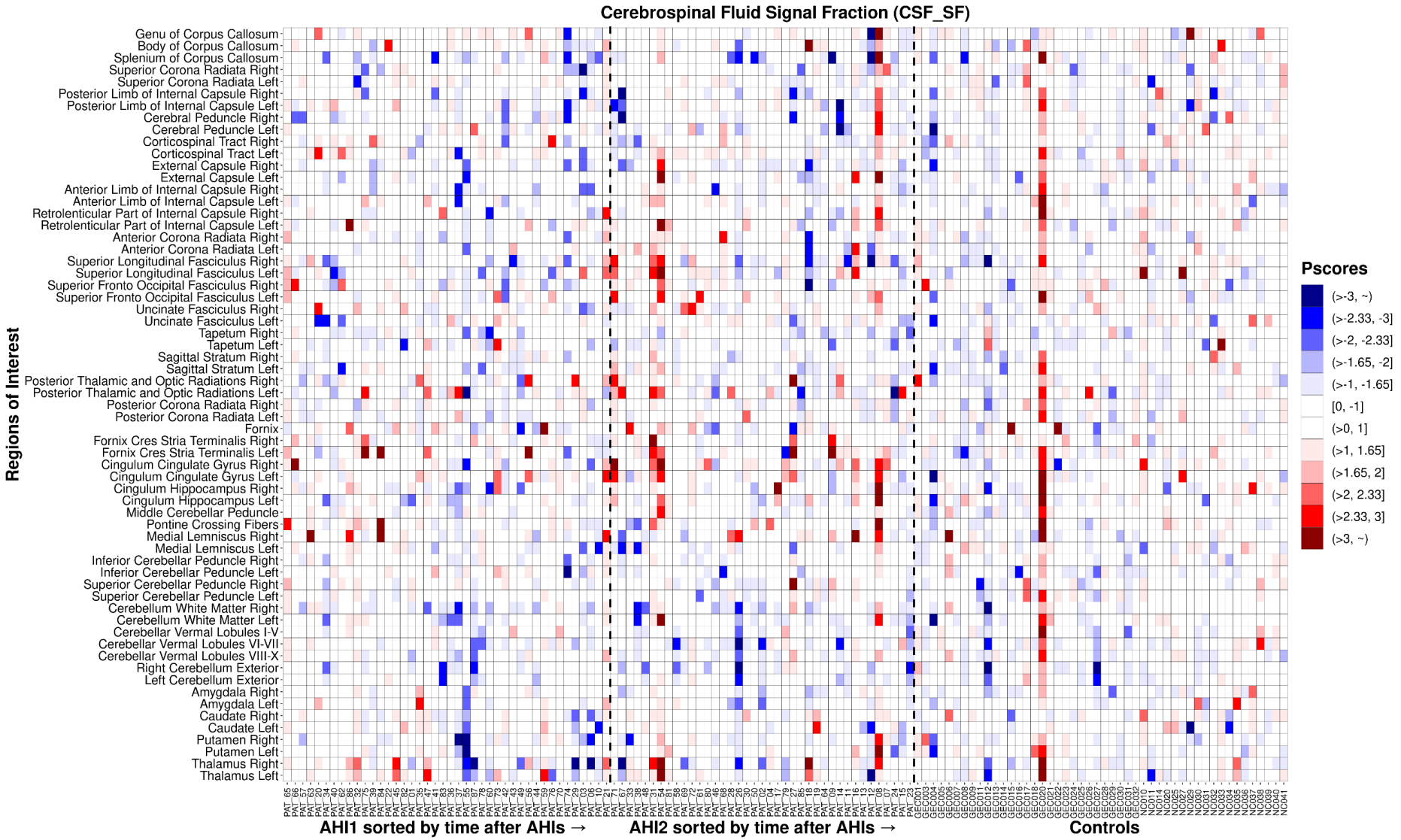
eFigure 7.I Heatmap of RTPP for participants with AHI and controls.



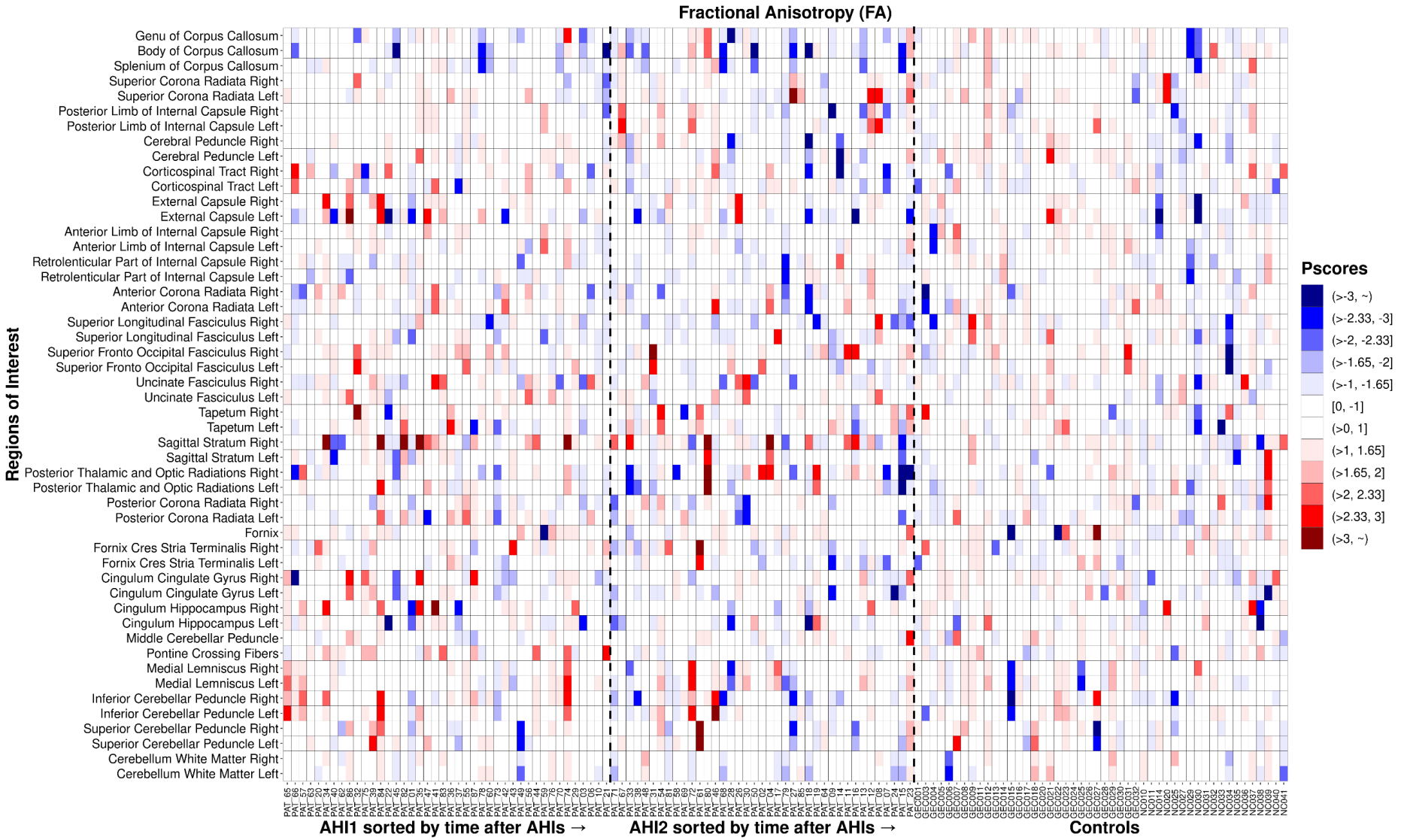
eFigure 8.A Heatmap of volumetrics for AHI1, AHI2 and control groups. The participants within each AHI category are sorted by increasing time (in days) after AHI. The global volumetric regions of interest are presented at the top (see in y-axis labels), followed by more local regions from the white matter (WM) and gray matter (GM). A few participants can be identified with extreme values (columns of red and blue tiles) but with no systematic patterns across ROIs (rows).



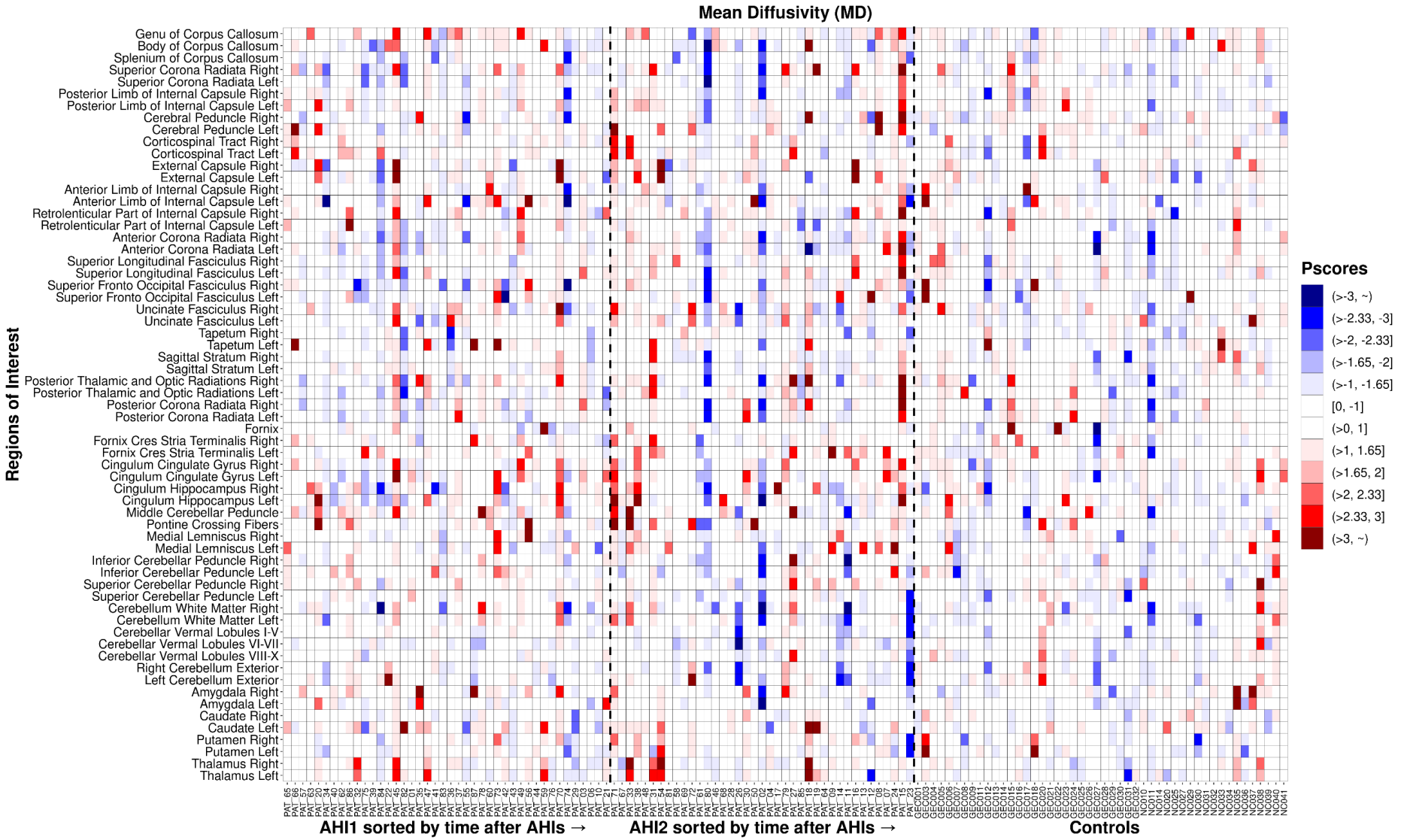
eFigure 8.B Heatmap of AD for AHI1, AHI2 and control groups. Few individuals with extreme values can be observed with no systematic patterns across ROIs (row). This format is consistent for the rest of the following figures from dMRI metrics.



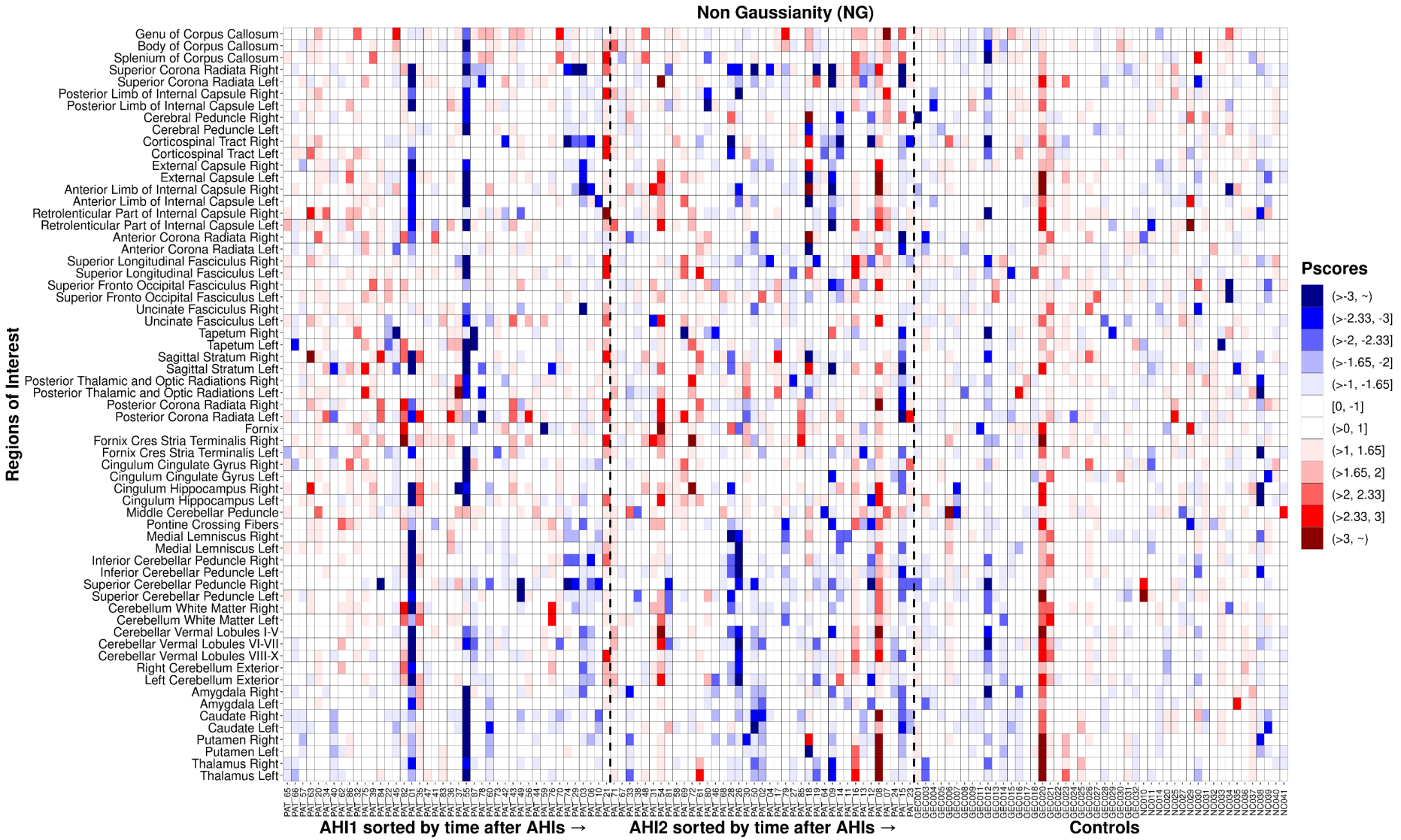
eFigure 8.C Heatmap of CSF-SF for AHI1, AHI2 and control groups.



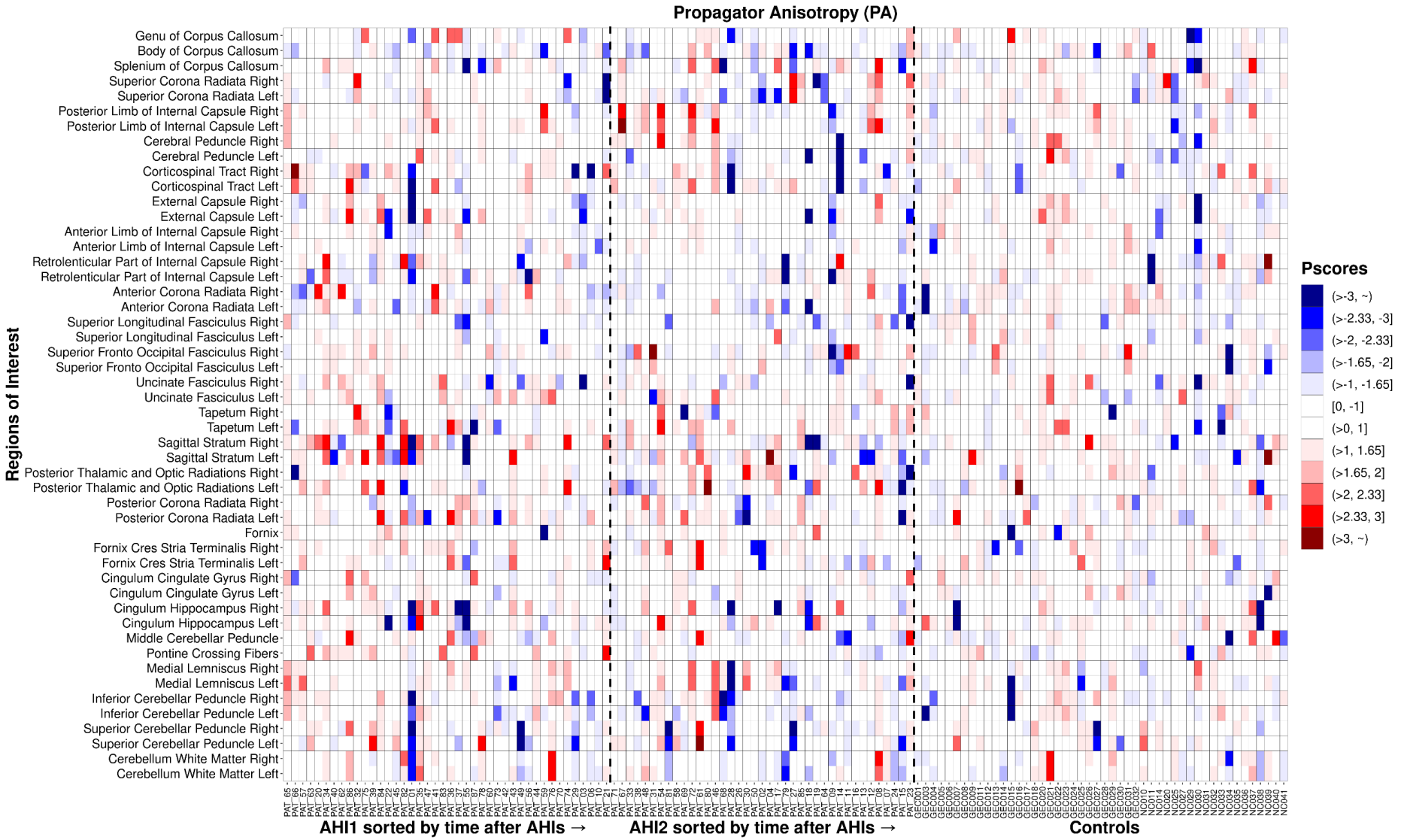
eFigure 8.D Heatmap of FA for AHI1, AHI2 and control groups.



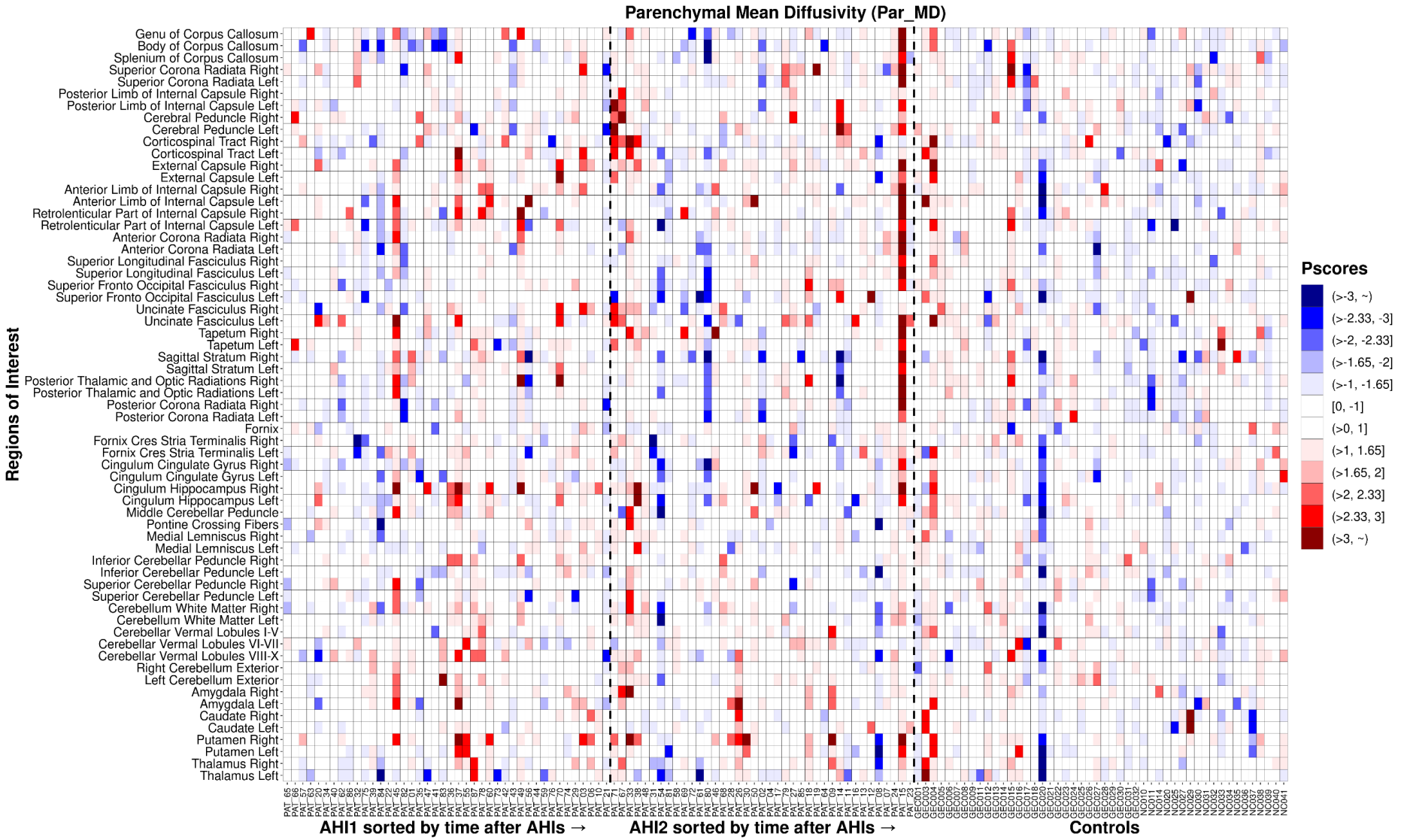
eFigure 8.E Heatmap of MD for AHI1, AHI2 and control groups.



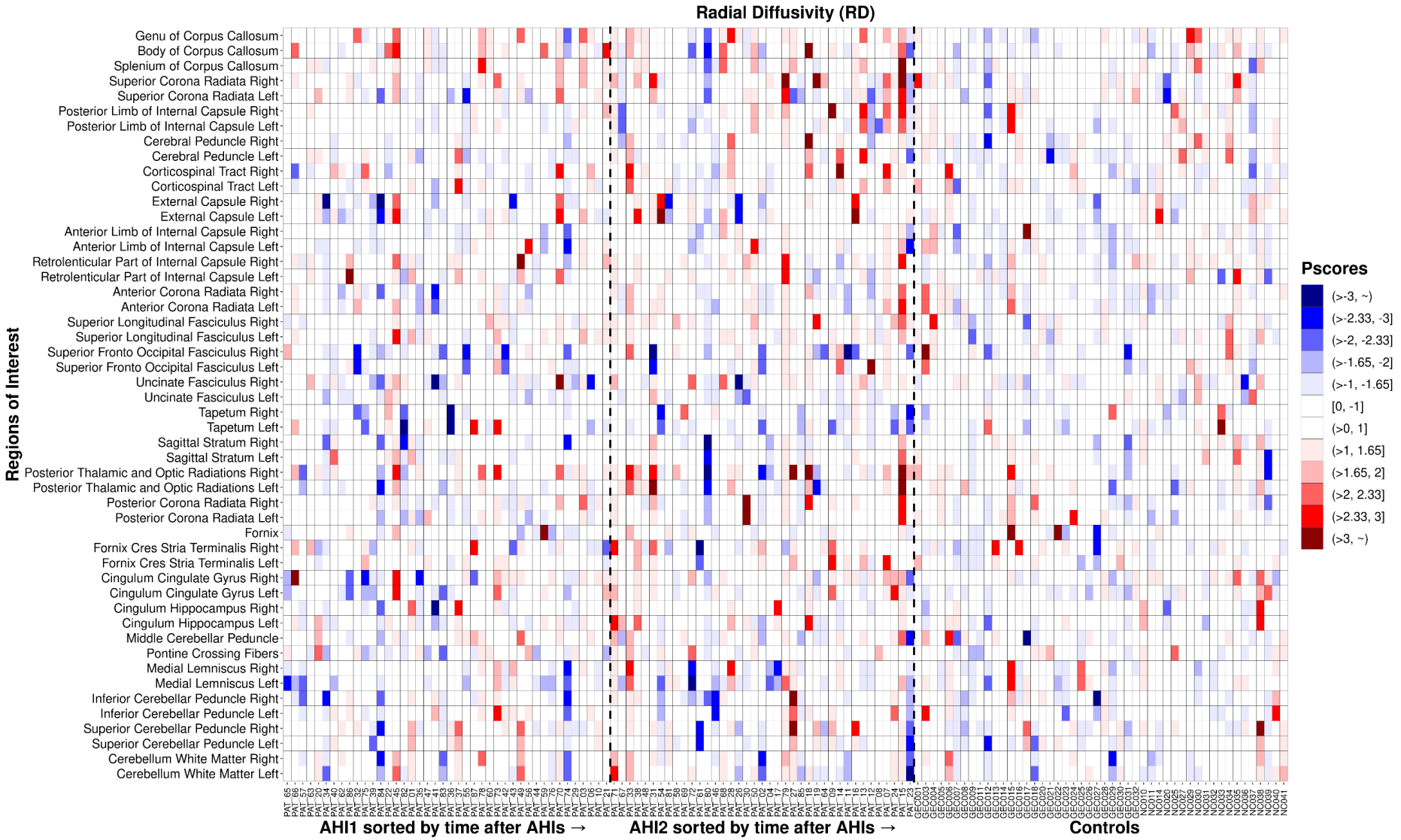
eFigure 8.F Heatmap of NG for AHI1, AHI2 and control groups.



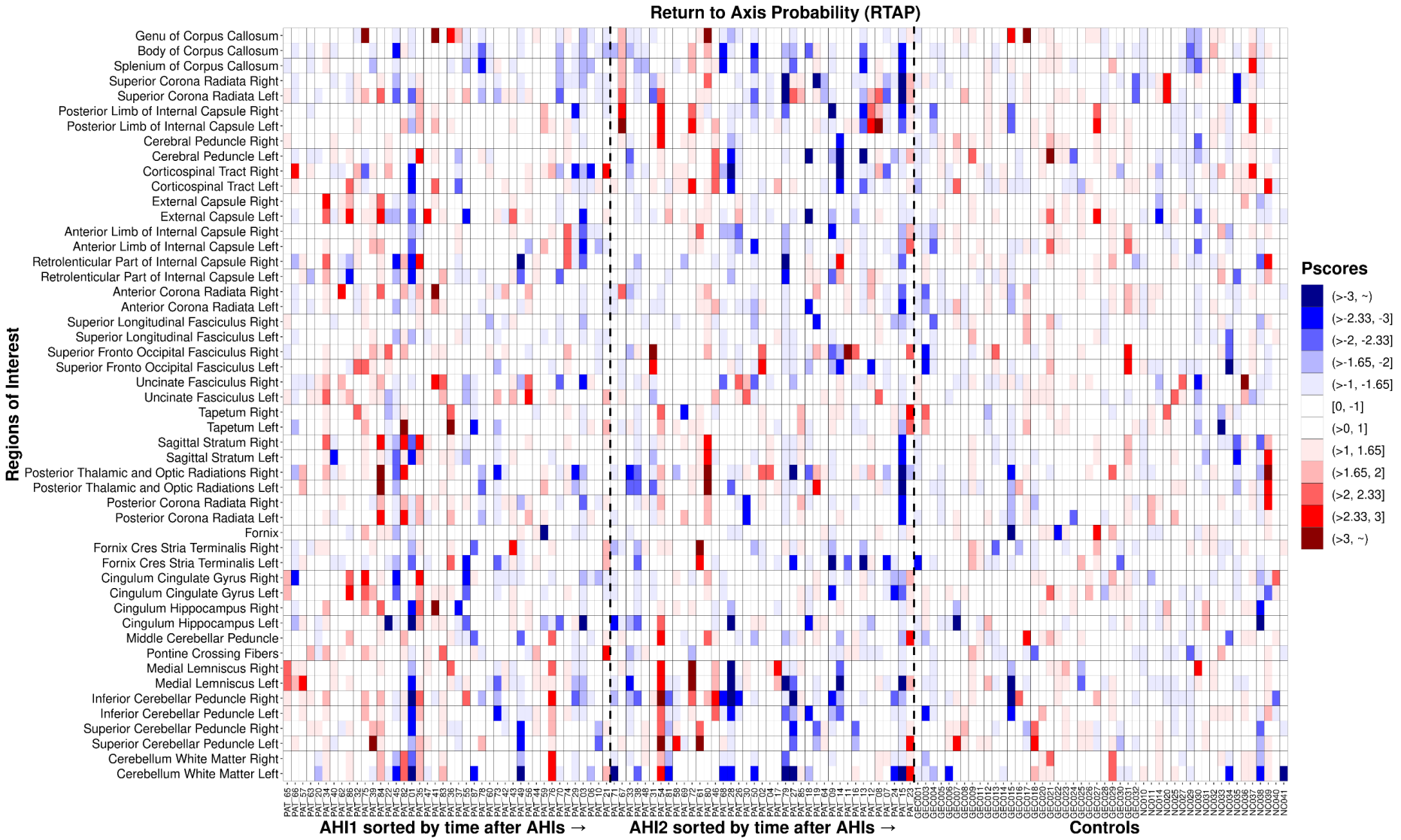
eFigure 8.G Heatmap of PA for AHI1, AHI2 and control groups.



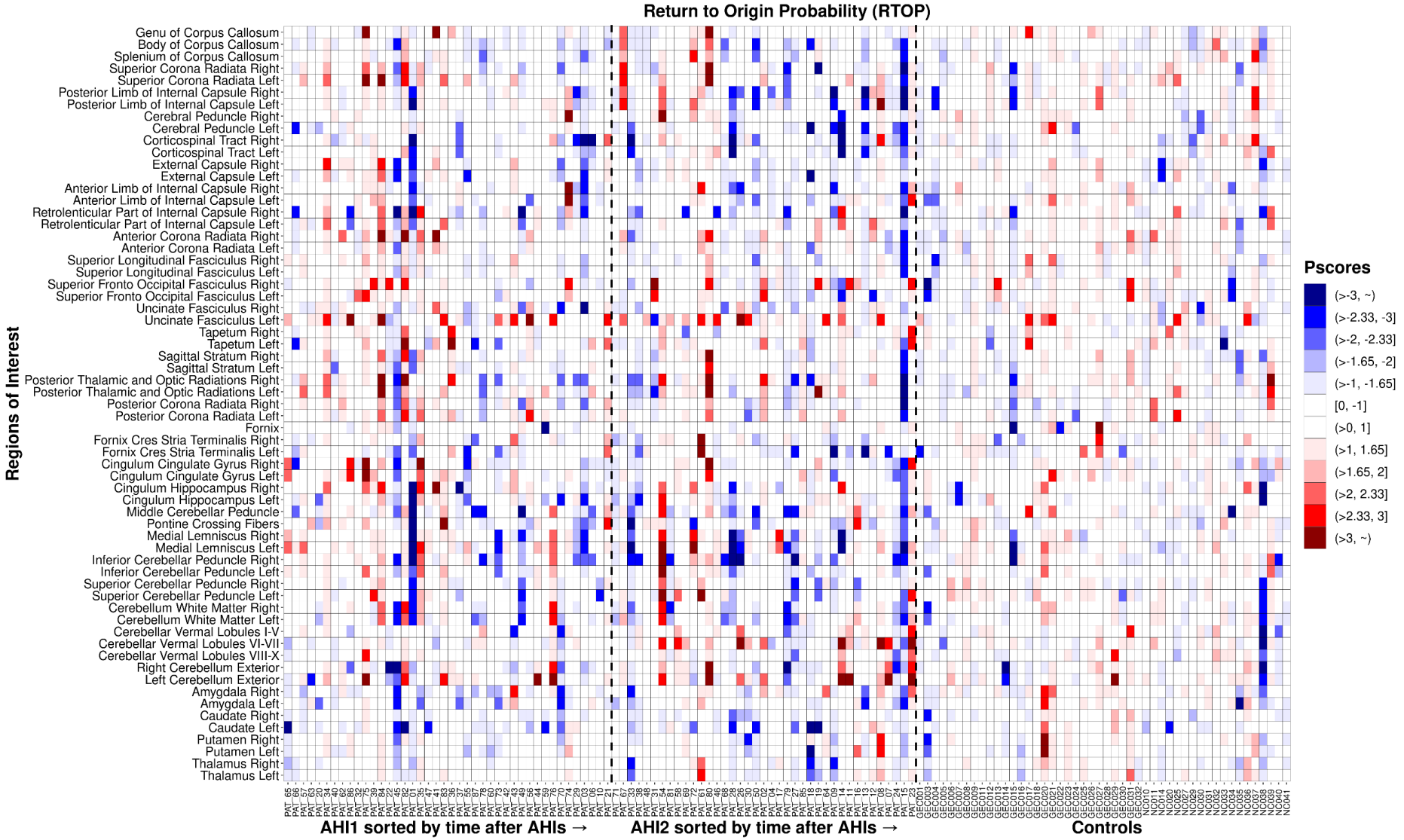
eFigure 8.H Heatmap of Par-MD for AHI1, AHI2 and control groups.



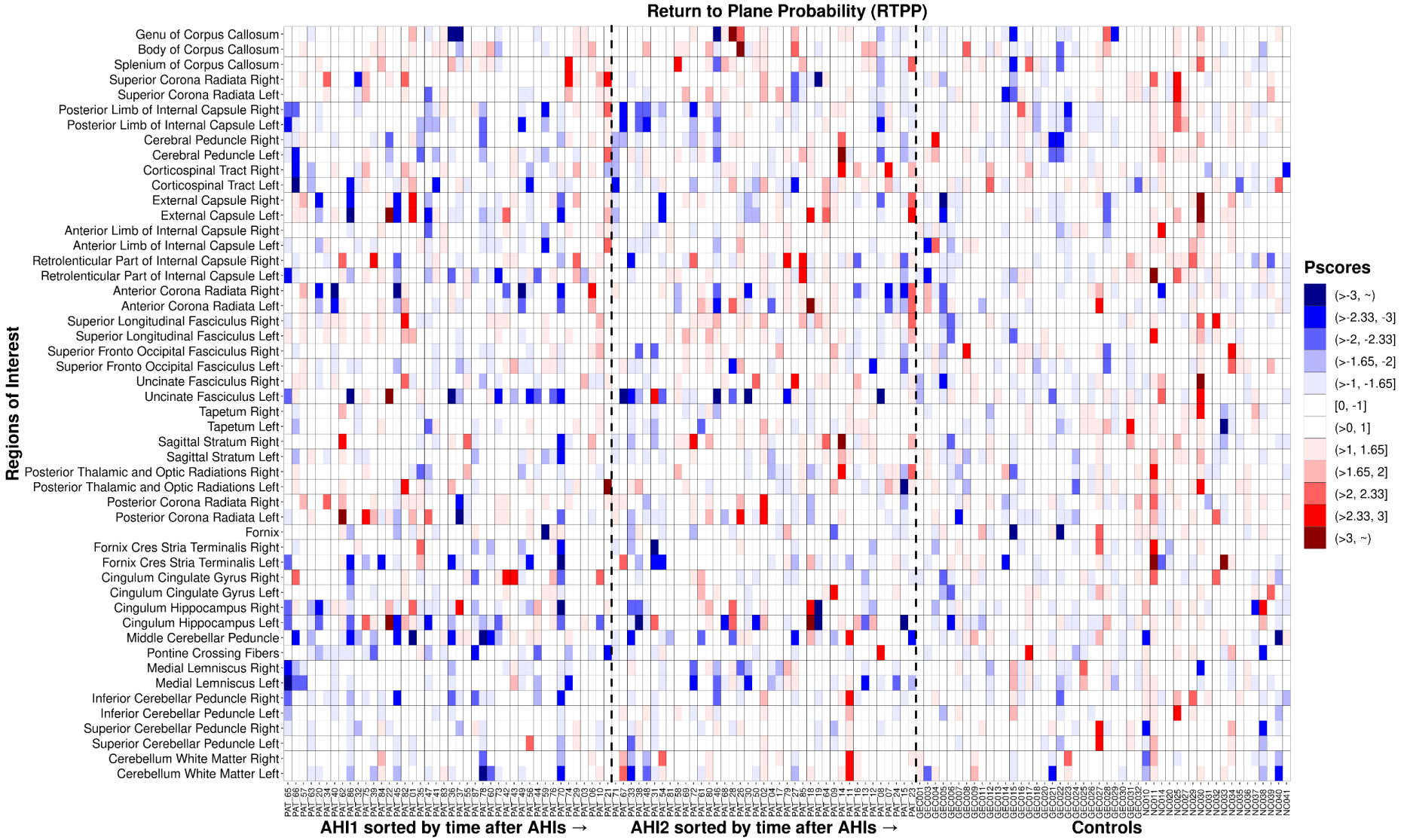
eFigure 8.I Heatmap of RD for AHI1, AHI2 and control groups.



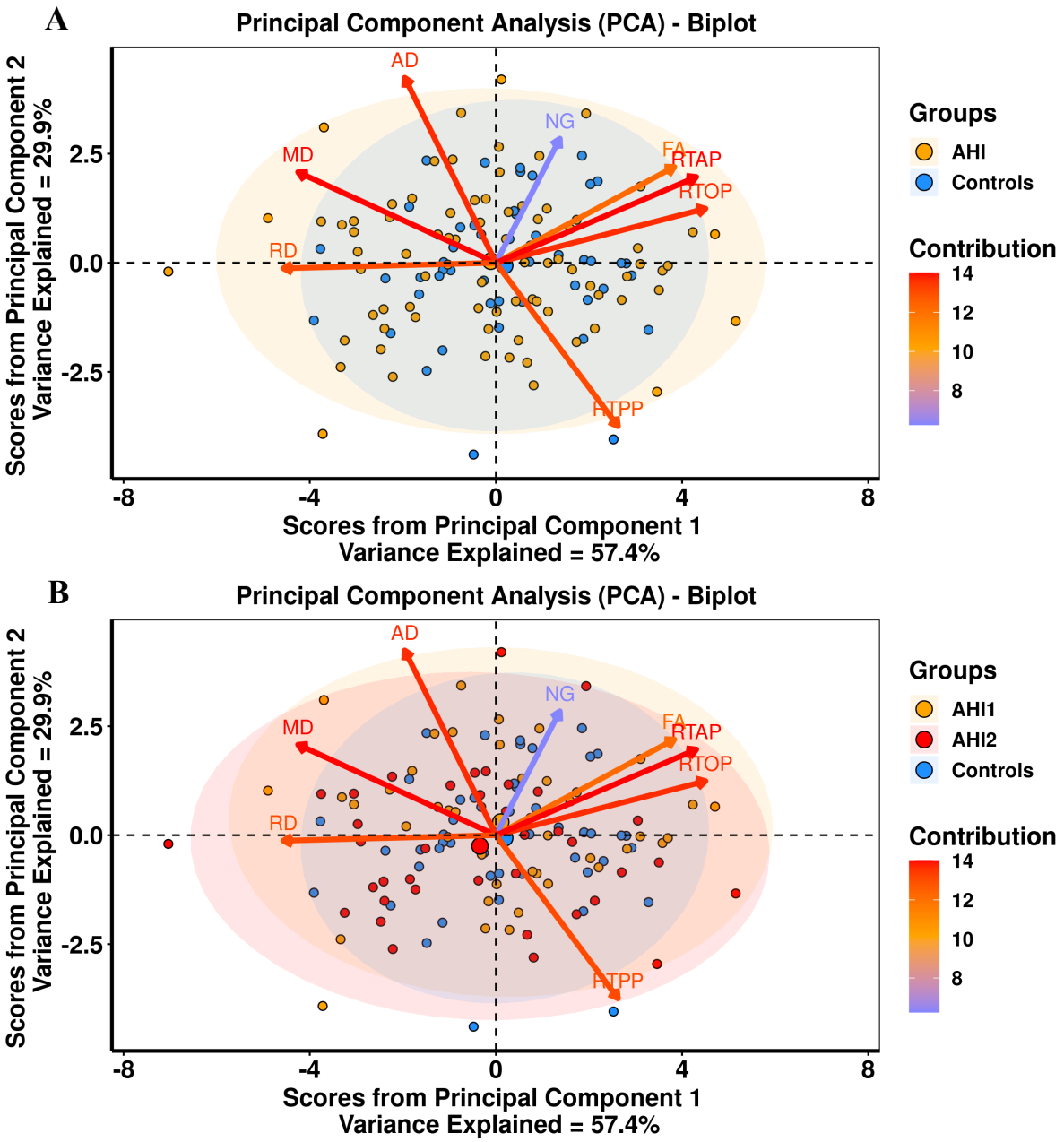
eFigure 8.J Heatmap of RTAP for AHI1, AHI2 and control groups.



eFigure 8.K Heatmap of RTOP for AHI1, AHI2 and control groups.

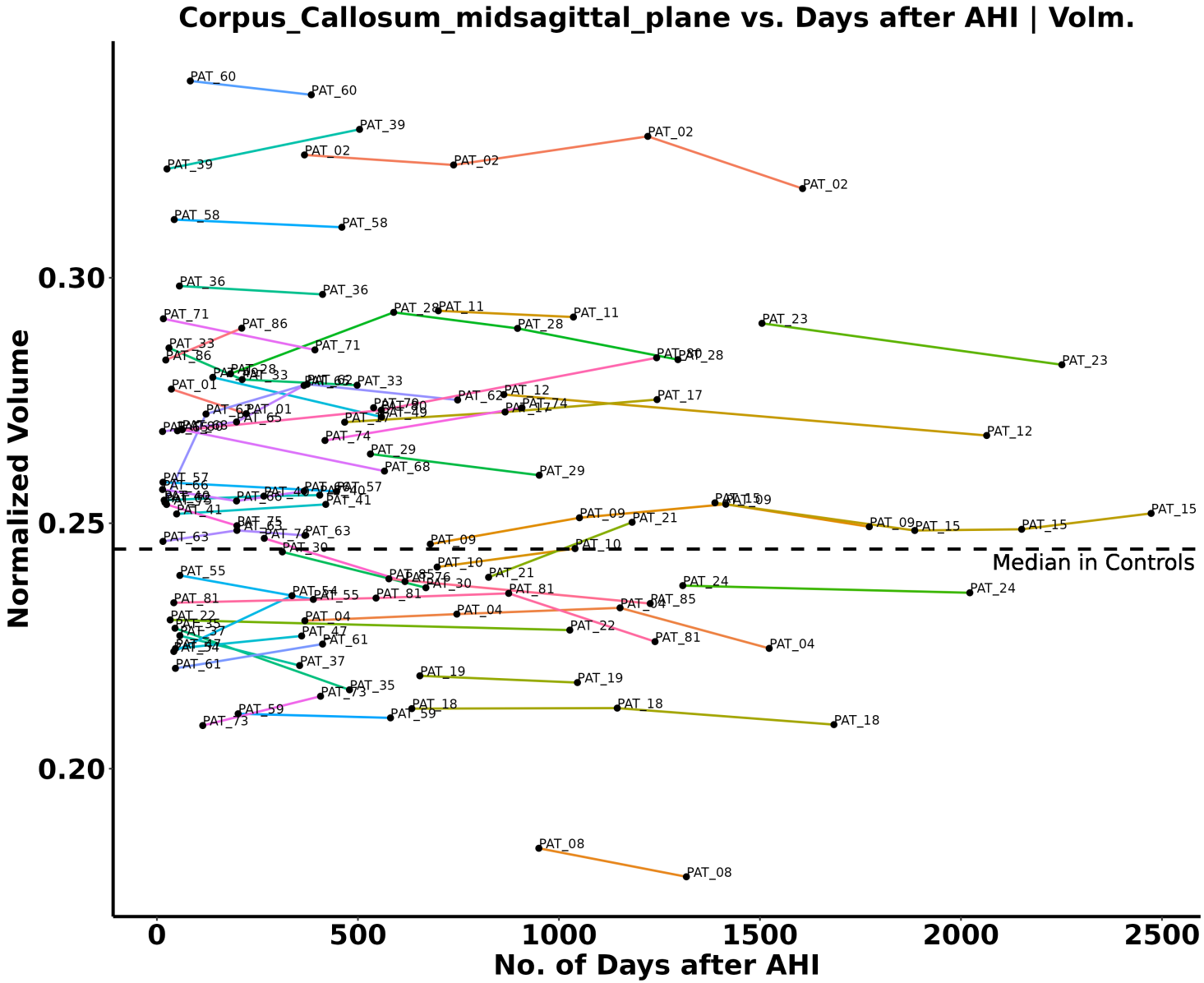


eFigure 8.L Heatmap of RTPP for AHI1, AHI2 and control groups.

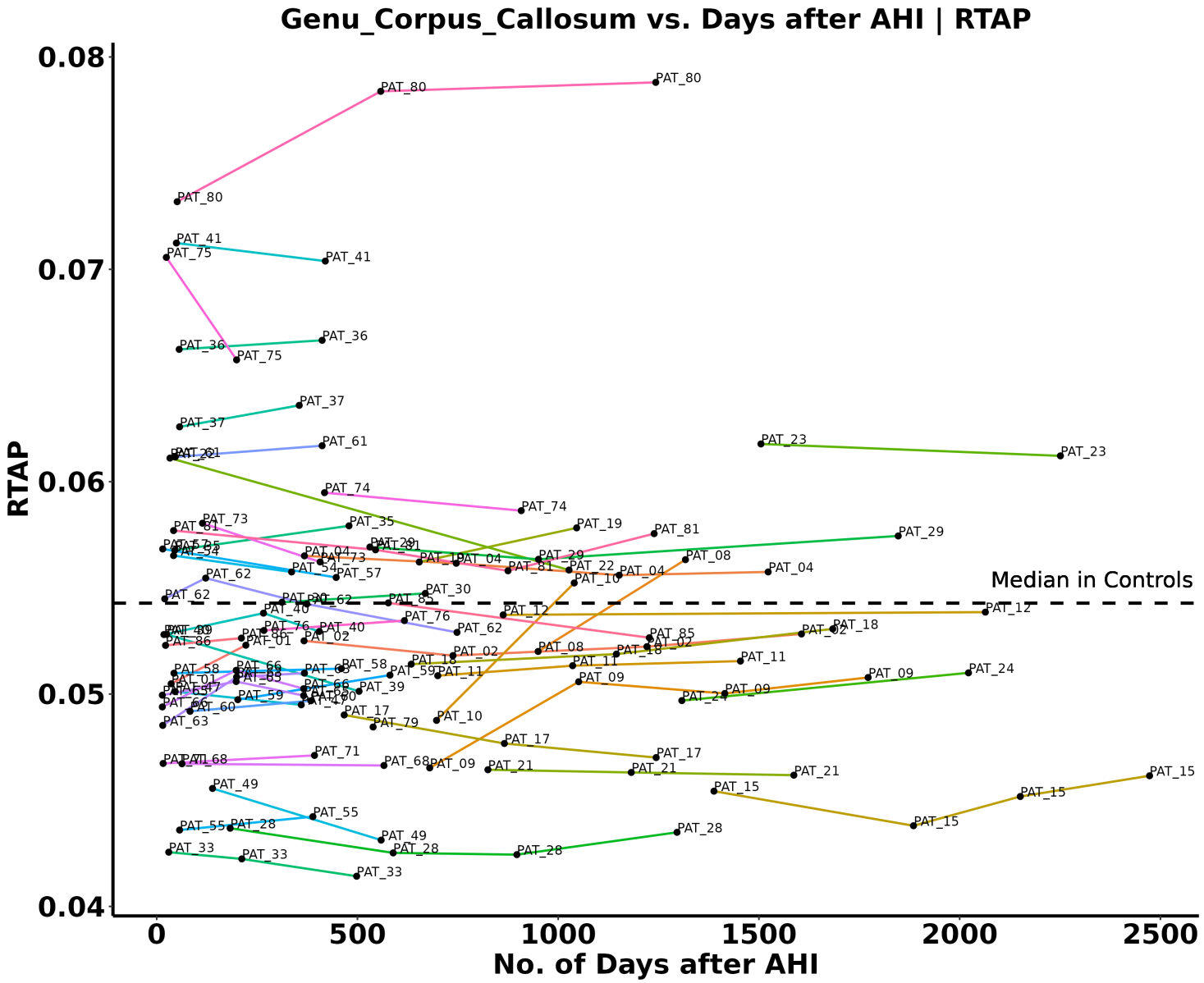


eFigure 9. Summary biplot of features from the first two principal components explaining highest variance in all combined white matter ROIs and eight examined diffusion MRI metrics. A (top) The x and y axes represent the individual participant scores for the first and the second principal components, respectively. The arrows represent the variables or diffusion MRI metrics. Please note, in biplots, the variables (arrows) are not in the same space as the individuals (dots). Therefore, it is recommended[71] to focus in the direction of the arrows rather than their absolute positions. The variables that are positively correlated (e.g., FA, RTAP, RTOP) appear together, while the variables that are negatively correlated appear on the opposite side of their respective quadrants (e.g., AD and RTPP). The length of the arrows informs about the representation of a variable to the first two principal components. Arrows that are further away from the origin and angled closer to the x-axis (e.g., RD, RTOP) are well represented in the first principal component, while those angled closer to the y-axis (AD, RTPP) are better represented in the second principal component. Arrows that are shorter and closer to the origin (e.g., NG) are not well represented by either of the first two principal components. The percentage (%) of contribution of each variable to

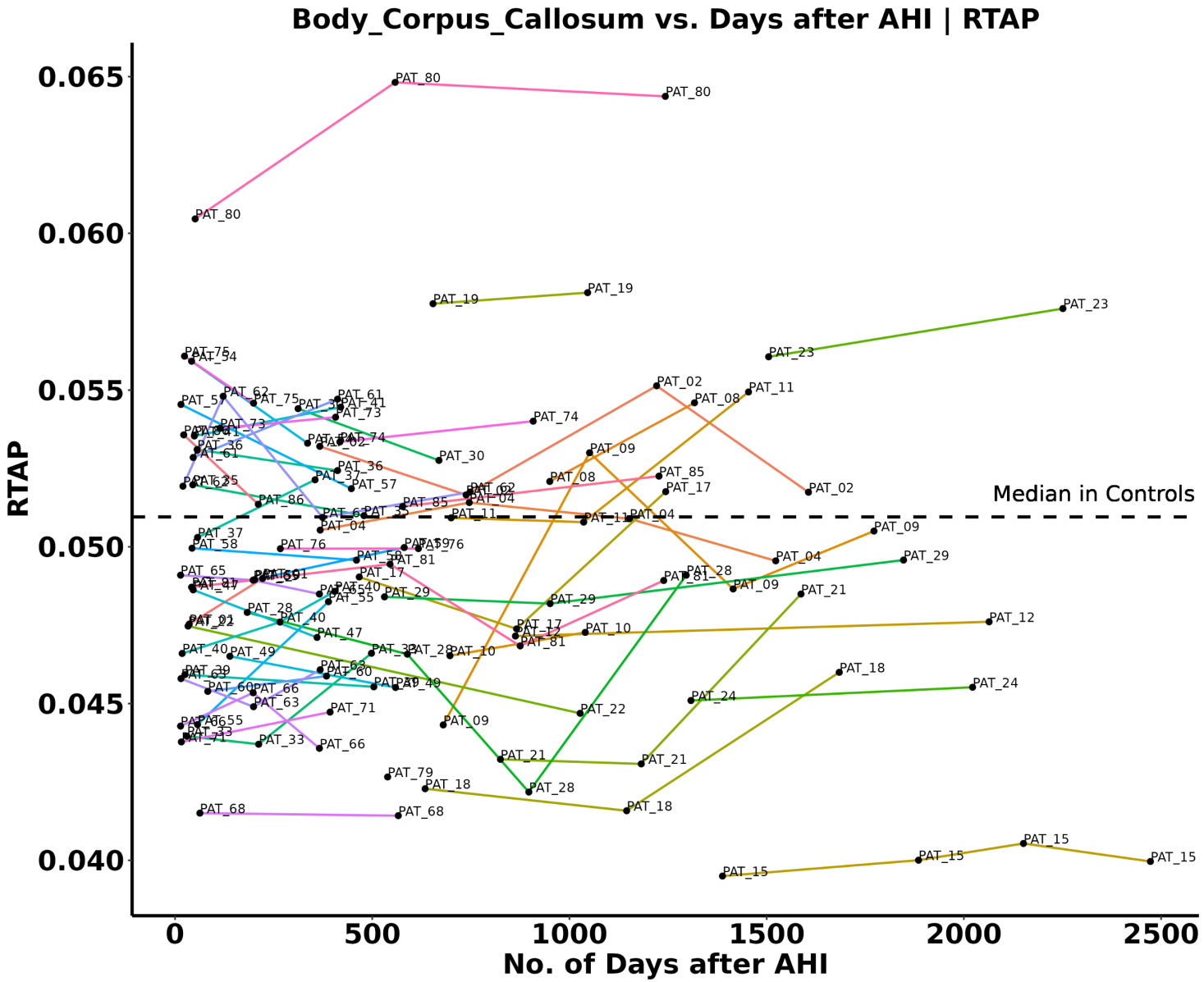
the first two principal components is represented by the gradient color map (bottom right legend) with higher contributions towards ‘red’. It indicates that the red arrows (diffusion metrics/variables) contribute more (> 12.5%) to the first two PCs compared to the expected percentage, if all eight variables contributed equally (100/8 = 12.5%). The individuals that appear on the same side of a variable (arrows) typically have large values for those variables, while those on the opposite side of the variable tend to have lower values. The orange and blue dots represent the observations from the AHI and Control groups, respectively, projected in the principal component analysis space. The larger dots represent the centers of the ellipses of each group, where each ellipse represents the 95% confidence interval of the corresponding group distributions. The confidence ellipse from the control group (light blue) completely overlaps with the ellipse from the AHI group (light orange) **B (bottom)** Same descriptions as above applies except incorporating the Control, AHI1 and AHI2 sub-groups. The orange, red and blue dots represent the observations from the AHI1, AHI2 and Control groups, respectively, projected in the principal component analysis space. The 95% confidence ellipses across all three groups largely overlap.



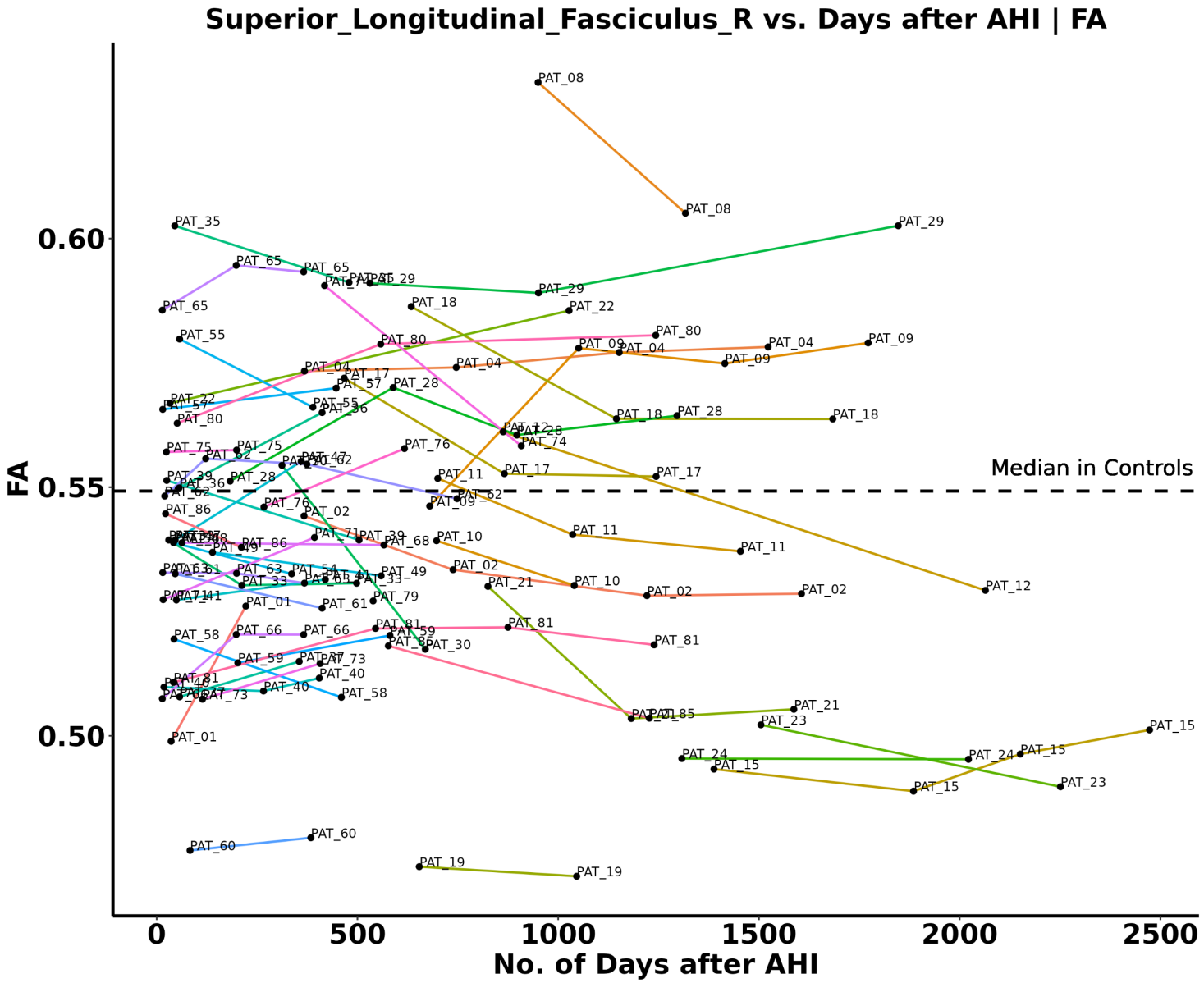
eFigure 10.A Longitudinal volumteric changes across time for participants with AHI within the Corpus Callosum Mid-Sagittal Plane. The y-axis represents normalized volumes and the x-axis represents time passed in days after AHI. The participants are dispersed in the y-axis in the baseline visit (dots shortly after 0 days). This explains the variability between subjects and group level differences in the cross-sectional analysis (compare with the dotted line for median in Controls). However, across time, remarkably small within subjects variability is observed. Each AHI participant ('PAT') label has a unique line color for their longitudinal trend. The number of repeated scans are the number of times a participant label appears with each successive dot across time.



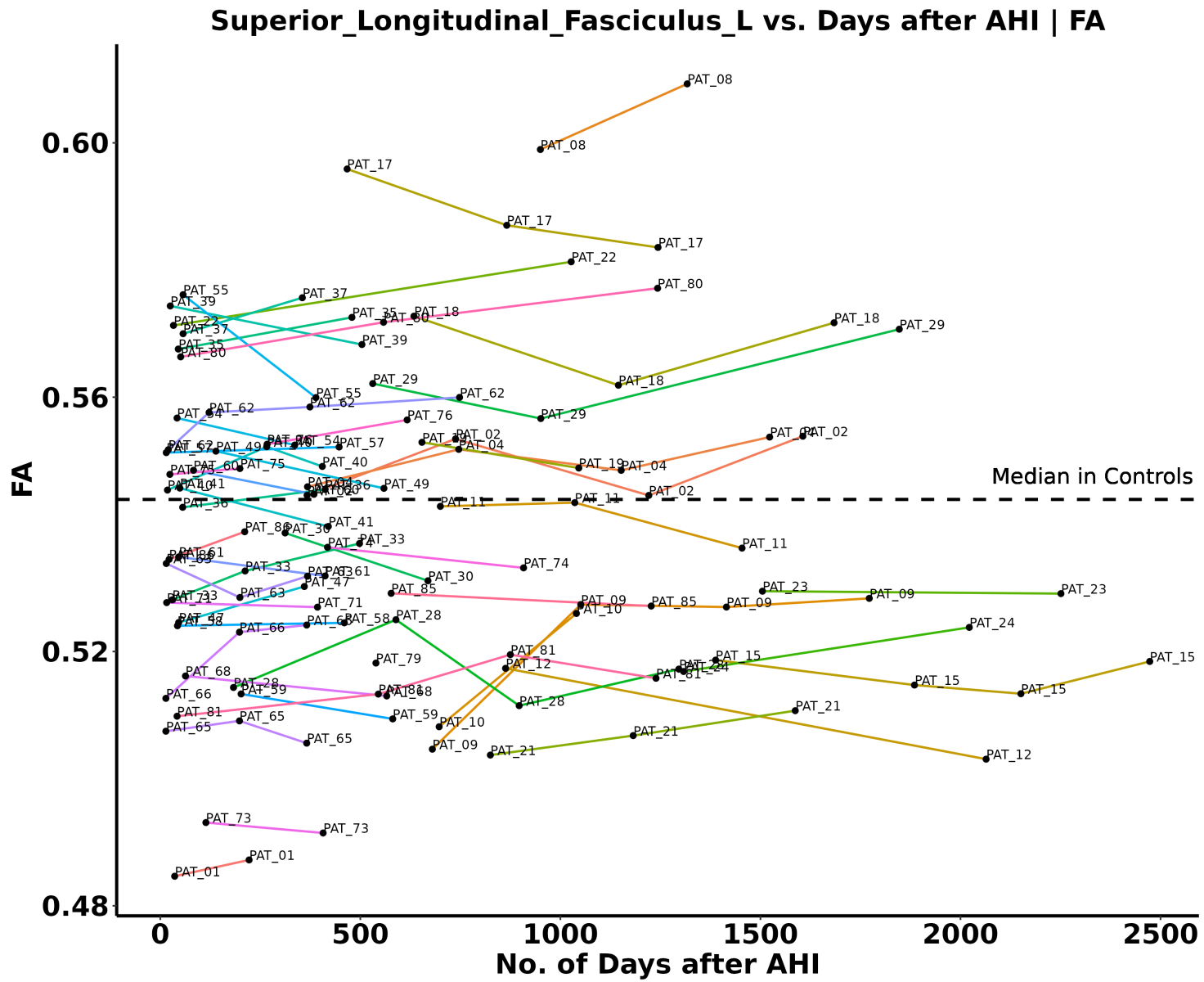
eFigure 10.B Longitudinal trends of diffusion changes across time in participants with AHI for RTAP within the Genu of the Corpus Callosum. The same descriptions as eFigure 10.A for volumetric data appllies to the diffusion data presented here. A very similar pattern of small within subjects variability and dispersed AHI participant distribution at the baseline visit can also be observed here.



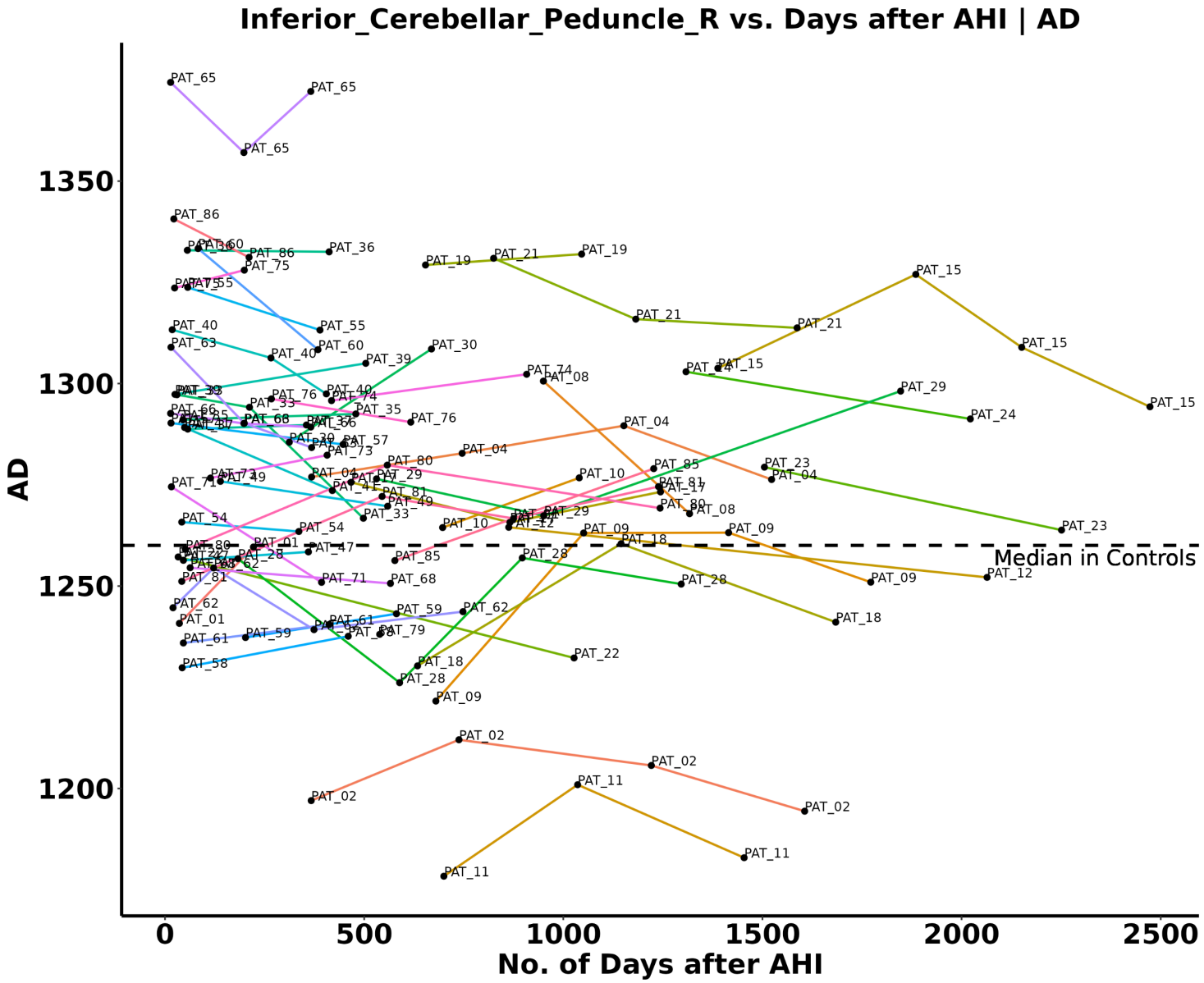
eFigure 10.C Similar descriptions as eFigure 10.B applies here, except within the Body of the Corpus Callosum. Similar to the Genu of the Corpus Callosum, most participants demonstrate consistent longitudinal trends of RTAP changes in this ROI.



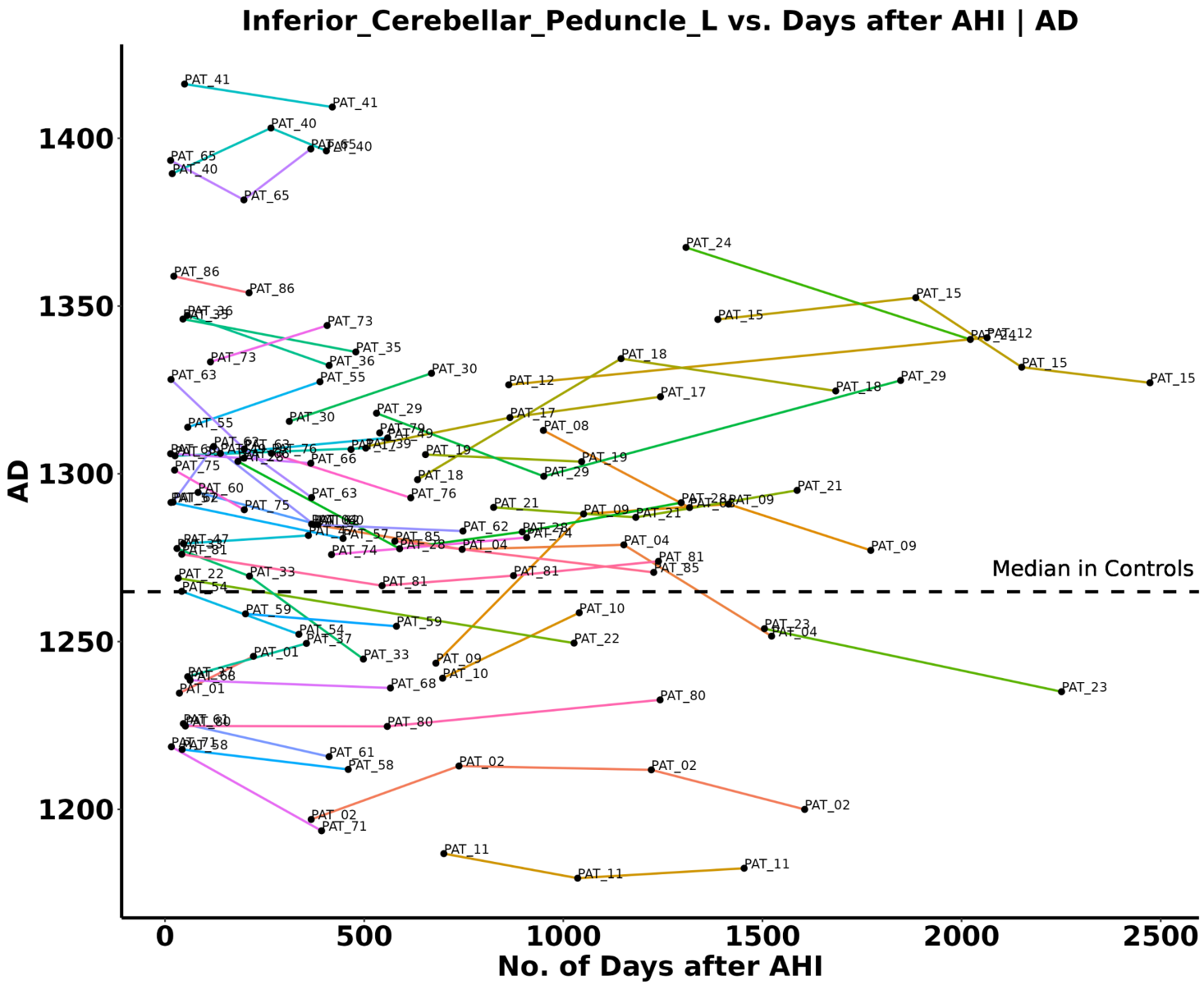
eFigure 10.D Similar descriptions as eFigure 10.A applies but for FA within the Right Superior Longitudinal Fasciculus. Like the other ROIs and metrics, it also demonstrates good stability against changes in FA across time for individual participants.



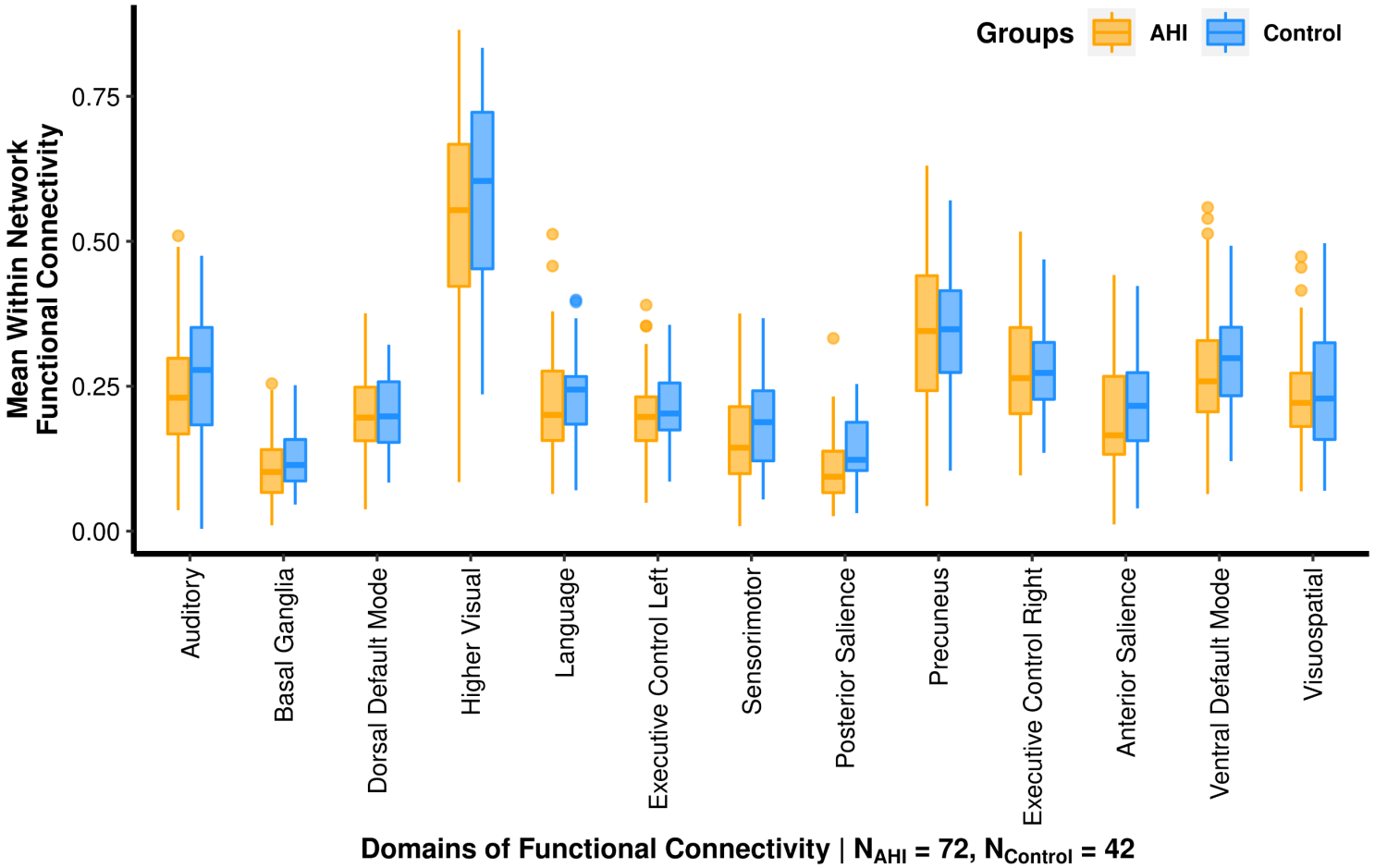
eFigure 10.E Essentially the bilateral counterpart of eFigure 10.D for FA. It shows similar stability as above but within the Left Superior Longitudinal Fasciculus.



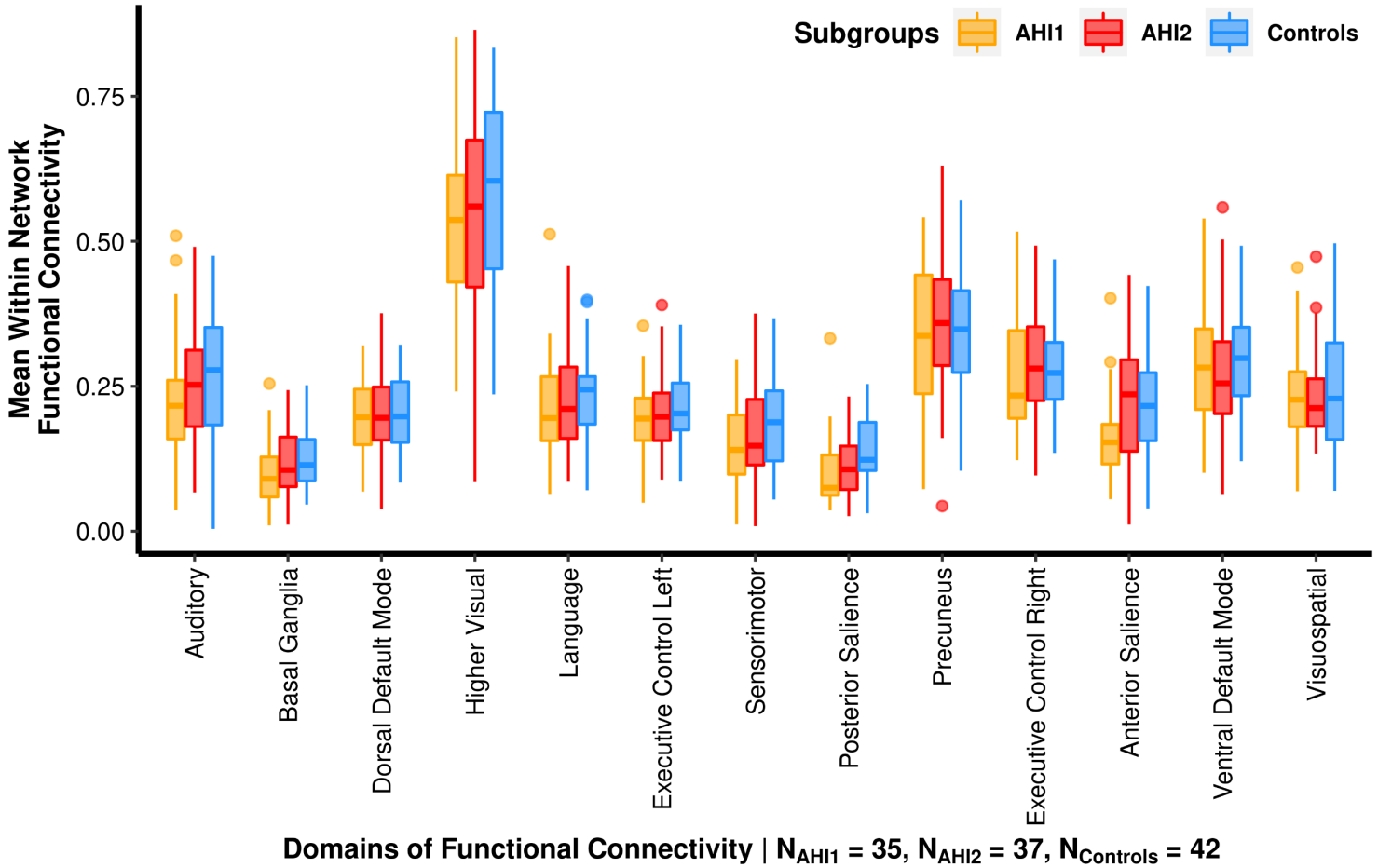
eFigure 10.F Similar descriptions as eFigure 10.A applies but for AD within the Right Inferior Cerebellar Peduncle. The participants demonstrate stable longitudinal changes in AD across time.



eFigure 10.G The same longitudinal stability can also be observed in the bilateral counterpart of eFigure 10.F for AD within the Left Inferior Cerebellar Peduncle.

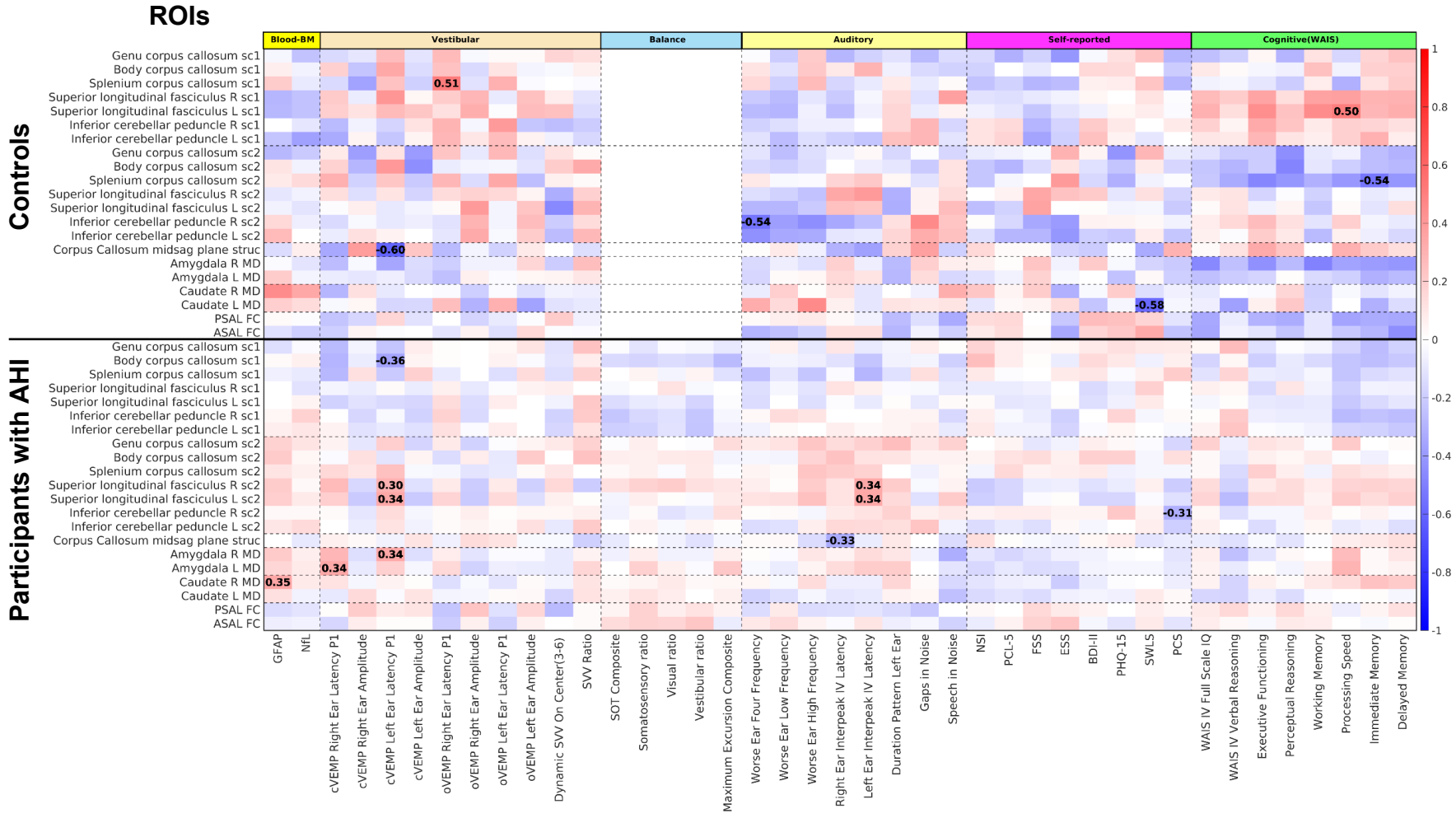


eFigure 11. Same as Figure 5 in the main text except repeated here to depict the Bayesian analysis results. The Posterior Salience (PSAL) network showed significant differences, where the participants with AHI demonstrated the weakest posterior probability to having enhanced within-network FC compared to Controls ($P^+ < 0.002$; i.e., < 0.2%). Similarly, more evidence was found, albeit with weaker posterior probabilities ($P^+ < 0.021$ and 0.042, i.e., < 2.1% and 4.2%, respectively) within the Sensorimotor (SMN) and Anterior Salience (ASAL) networks, respectively.



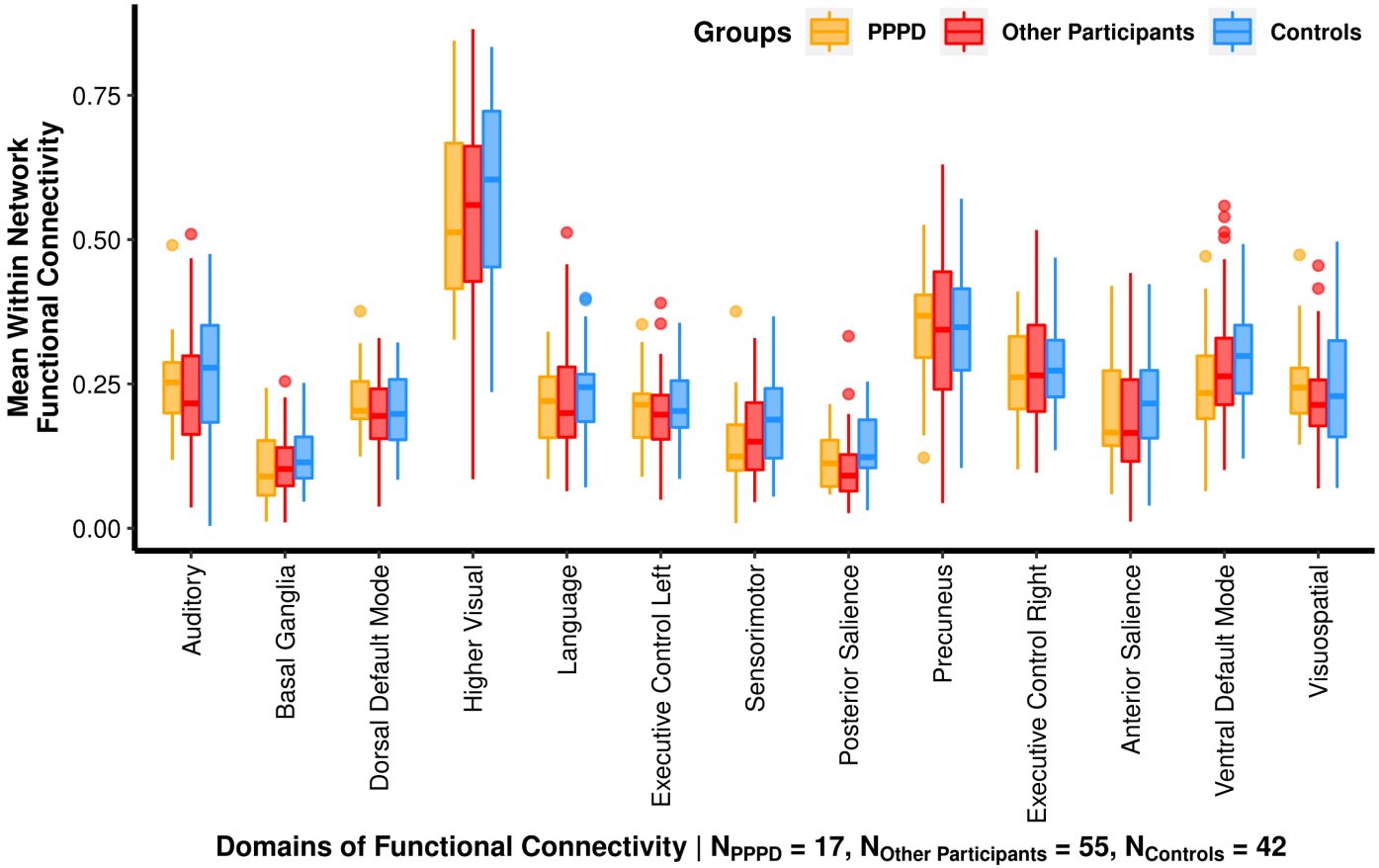
eFigure 12. Box plots showing within network functional connectivity across Controls, AHI1 and AHI2 participants. The y-axis represents the within-network FC values, and the x-axis represents the resting state networks. The orange, red and blue bars represent AHI1, AHI2 and Control groups, respectively. The horizontal lines represent the medians in each group. The vertical lines represent the spread of FC values within each group up to 1.5 times the inter-quartile range (IQR). The dots represent more extreme values. The figure indicates strong evidence towards AHI1 participants demonstrating less FC within the anterior and posterior salience networks (ASAL and PSAL) compared to Controls (Benjamini-Hochberg adjusted **P_{BH} = 0.021 and 0.029, respectively). We observed similar outcomes from the Bayesian analysis for the ASAL and PSAL

networks, where the AHI1 group demonstrated the weakest posterior probability to having enhanced within-network FC compared to Controls ($P^+ < 0.002$; i.e., < 0.2%, respectively). More evidence to AHI1 demonstrating weaker posterior probabilities ($P^+ < 0.026$ and 0.028, i.e., < 2.6% and 2.8%, respectively) to having enhanced FC compared to Controls in the SMN and BG networks were also observed. Therefore, both the conventional and Bayesian approaches provide the strongest evidence for AHI1 having less within-network FC compared to Controls in the ASAL and PSAL networks.

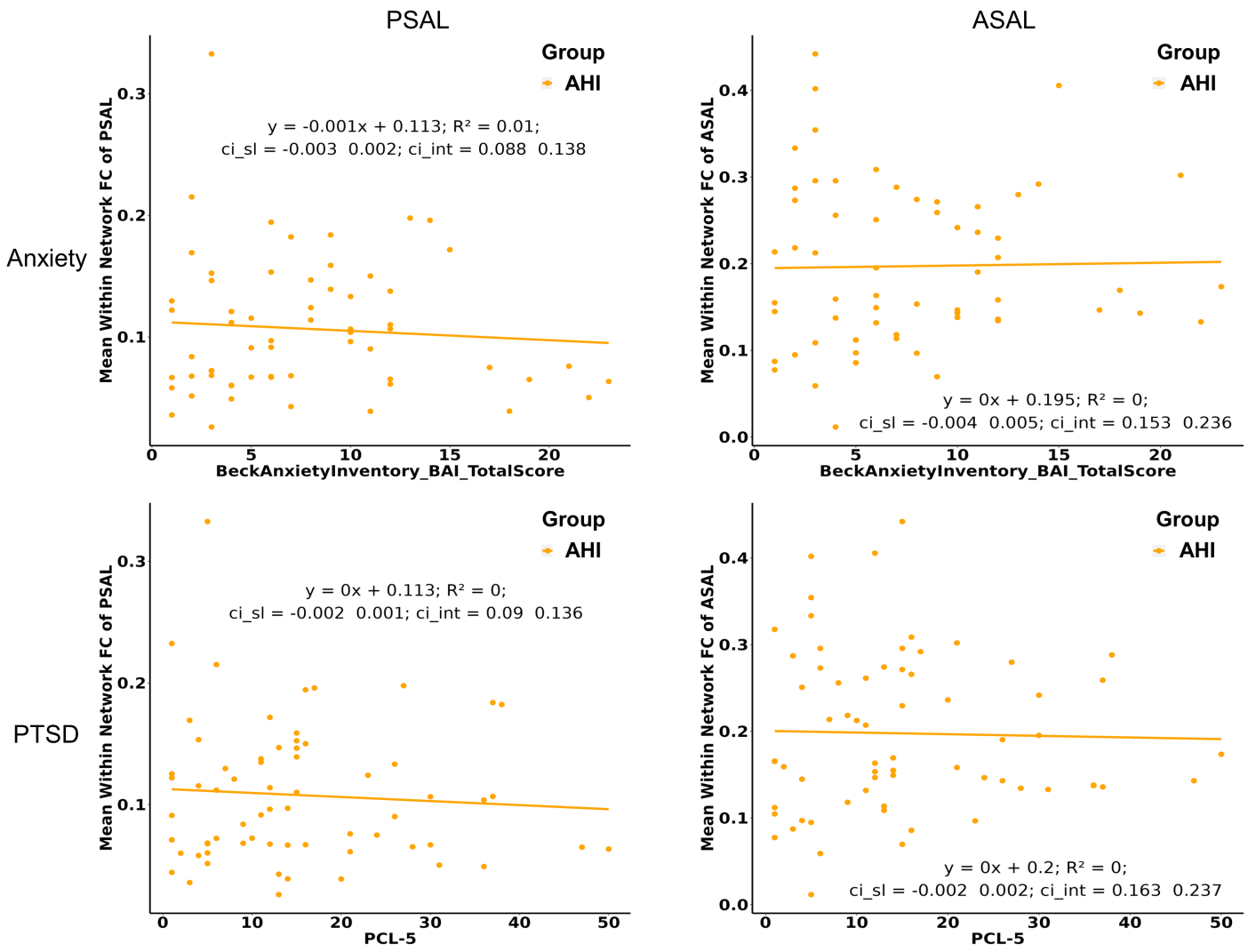


eFigure 13. Analysis of the correlation between imaging and clinical measures for ROIs with imaging metrics that were found significantly different between participants and controls at the group level. The figure is divided by a line that separates the correlation values from the same ROIs in the control (top) and participants with AHI (bottom) groups. The ROI names and the corresponding metric are provided on the left side of the y-axis. The first seven rows (top) show the correlation of the PCA scores obtained from diffusion imaging within white matter ROIs for

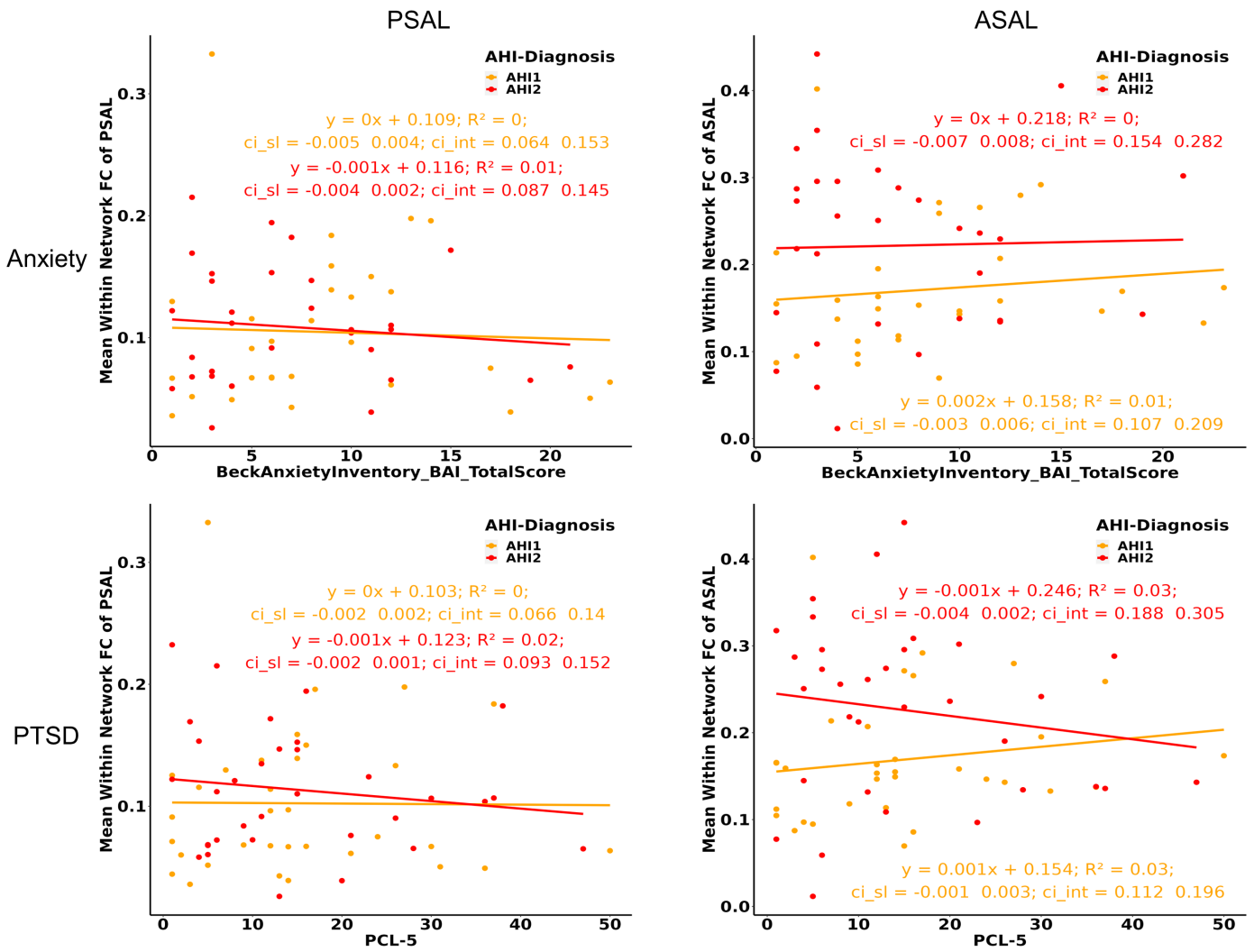
the 1st component ('sc1') and the next seven rows (top) show the correlations of PCA scores from the 2nd component (sc2). The 15th row (top) shows the correlations for the structural volumetric measurement within the Corpus Callosum mid-sagittal plane. Rows 16-19 show the correlations of MD within gray matter ROIs. The last two rows (top) show the correlations of FC from the salience networks (PSAL FC and ASAL FC). The same order is repeated for participants below the dividing line. The x-axis shows all 41 clinical parameters which were correlated to the ROI metrics in the y-axis. The principal domains from which corresponding clinical parameters originate are shown with a color code on the top of the correlation matrix. The color bar on the right shows the range of correlation values and associated colors, with blue representing negative correlations and red representing positive correlation values. Specific correlation values shown in the plot are *Spearman's rho (ρ)* values for those metrics that demonstrate a significant correlation at $P < 0.01$. The dashed lines help separate the correlations for specific ROI metrics and clinical domain for visual aesthetics. The blank white area under "Balance" (top) represents $\rho = 0$ values, as these data were not available from the control group. Abbreviations*: sc1* = scores from the 1st principal component; *sc2* = scores from the 2nd principal component; *R* = right; *L* = left; *struc* = structural volumetric measurement; *PSAL FC* = Mean Functional Connectivity within the Posterior Salience network; *ASAL FC* = Mean Functional Connectivity within the Anterior Salience network; *Blood-BM* = Blood Biomarker; *WAIS* = Wechler Adult Intelligence Scale; *GFAP* = Glial Fibrillary Acidic Protein; *NfL* = Neurofilament light chain; *cVEMP* = cervical vestibular evoked myogenic potential; *oVEMP* = ocular vestibular evoked myogenic potential; *SVV* = subjective visual vertical; *SOT* = Sensory Organization Test; *NSI* = Neurobehavioral Symptoms Inventory; *PCL-5* = PTSD-Check List version 5; *FSS* = Fatigue Severity Scale; *ESS* = Epworth Sleepiness Scale; *BDI-II* = Beck's Depression Inventory II; *PHQ-15* = Patient Health Questionnaire-15; *SWLS* = Satisfaction with Life Scale; *PCS* = Pain Catastrophizing Scale.



eFigure 14. Box plots showing within network functional connectivity across Controls, participants with PPPD and other participants without PPPD ('Other Participants'). The y and x axes represent the within-network FC and the resting state networks, respectively. The orange, red and blue bars represent the 'PPPD', 'Other Participants' and 'Controls', respectively. The horizontal lines represent the medians in each group. The vertical lines represent the spread of functional connectivity values within each group up to 1.5 times the inter-quartile range. The dots represent more extreme values. From an omnibus Kruskal-Wallis test across the three groups at an unadjusted level ($P_{KW} = 0.009$), significant difference was observed for the Posterior Salience (PSAL) network from a *post-hoc* Mann-Whitney *U* test adjusted for multiple comparisons, $P_{BH} < 0.05$. However, this evidence was towards the 'Other Participants' demonstrating less FC compared to Controls (**$P_{BH} = 0.009$). The PPPD group showed no significant difference in FC with the Controls for any of the networks.



eFigure 15.A Participants with AHI show no significant relationship of anxiety and PTSD measures with FC in the salience networks. The top row shows the relationship between anxiety and FC within the PSAL and ASAL networks. The x-axis represents the Beck Anxiety Inventory (BAI) Total Scores, and the y-axis represents the FC within corresponding networks, respectively. Similarly, the bottom row shows the relationships between PTSD (PCL-5) and FC in the PSAL and ASAL networks. The equation for each plot is given within the figure, along with the R^2 values. The confidence intervals for the slope ('ci_sl') and intercept ('ci_int') are also provided in each figure.



eFigure 15.B AHI subgroups (AHI1 and AHI2) show no significant relationship of anxiety and PTSD measures with FC in the salience networks. The same descriptions as eFigure 15.A applies here, except for the two AHI subgroups separately, orange – for AHI1 and red – for AHI2.

References

1. Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *NeuroImage*. Feb 1999;9(2):179-94. doi:10.1006/nimg.1998.0395
2. Roy S, Butman JA, Pham DL, Alzheimers Disease Neuroimaging I. Robust skull stripping using multiple MR image contrasts insensitive to pathology. *Neuroimage*. Feb 1 2017;146:132-147. doi:10.1016/j.neuroimage.2016.11.017
3. Lao Z, Shen D, Xue Z, Karacali B, Resnick SM, Davatzikos C. Morphological classification of brains via high-dimensional shape transformations and machine learning methods. *Neuroimage*. Jan 2004;21(1):46-57. doi:10.1016/j.neuroimage.2003.09.027
4. Davatzikos C, Genc A, Xu D, Resnick SM. Voxel-based morphometry using the RAVENS maps: methods and validation using simulated longitudinal atrophy. *NeuroImage*. Dec 2001;14(6):1361-9. doi:10.1006/nimg.2001.0937
5. Goldszal AF, Davatzikos C, Pham DL, Yan M, Bryan RN, Resnick SM. An image-processing system for qualitative and quantitative volumetric analysis of brain images. *Journal of computer assisted tomography*. 1998;22 5:827-37. doi:10.1097/00004728-199809000-00030
6. Doshi J, Erus G, Ou Y, et al. MUSE: MUlti-atlas region Segmentation utilizing Ensembles of registration algorithms and parameters, and locally optimal atlas selection. *NeuroImage*. Feb 15 2016;127:186-195. doi:10.1016/j.neuroimage.2015.11.073
7. Oishi K, Zilles K, Amunts K, et al. Human brain white matter atlas: identification and assignment of common anatomical structures in superficial white matter. *NeuroImage*. 2008;43(3):447-457. doi:10.1016/j.neuroimage.2008.07.009
8. Van Essen DC, Ugurbil K, Auerbach E, et al. The Human Connectome Project: a data acquisition perspective. *NeuroImage*. Oct 1 2012;62(4):2222-31. doi:10.1016/j.neuroimage.2012.02.018
9. Tustison NJ, Avants BB, Cook PA, et al. N4ITK: improved N3 bias correction. *IEEE Trans Med Imaging*. Jun 2010;29(6):1310-20. doi:10.1109/tmi.2010.2046908
10. Klasson N OE, Eckerström C, Malmgren H, Wallin A. Estimated intracranial volume from FreeSurfer is biased by total brain volume. *Eur Radiol Exp 2018 Sep 19;2:24 doi: 101186/s41747-018-0055-4 PMCID: PMC6143491*.
11. Verma R, Swanson RL, Parker D, et al. Neuroimaging Findings in US Government Personnel With Possible Exposure to Directional Phenomena in Havana, Cuba. *JAMA*. 2019;322(4):336-347. doi:10.1001/jama.2019.9269
12. Ou Y, Sotiras A, Paragios N, Davatzikos C. DRAMMS: Deformable registration via attribute matching and mutual-saliency weighting. *Med Image Anal*. Aug 2011;15(4):622-39. doi:10.1016/j.media.2010.07.002
13. Oishi K, Faria A, Jiang H, et al. Atlas-based whole brain white matter analysis using large deformation diffeomorphic metric mapping: application to normal elderly and Alzheimer's disease participants. *NeuroImage*. Jun 2009;46(2):486-99. doi:10.1016/j.neuroimage.2009.01.002
14. Oishi K, Faria A, Mori S. JHU-MNI-ss Atlas. https://www.slicer.org/publications/item/view/1883; Johns Hopkins University School of Medicine, Department of Radiology, Center for Brain Imaging Science.
15. Irfanoglu MO, Sarlls J, Nayak A, Pierpaoli C. Evaluating corrections for Eddy-currents and other EPI distortions in diffusion MRI: methodology and a dataset for benchmarking. *Magnetic resonance in medicine*. Apr 2019;81(4):2774-2787. doi:10.1002/mrm.27577
16. Irfanoglu MO, Nayak A, Jenkins J, Pierpaoli C. TORTOISE v3: Improvements and New Features of the NIH Diffusion MRI Processing Pipeline. ISMRM, 2018. 2018:

17. Pierpaoli C, Lindsay W, Irfanoglu MO, et al. TORTOISE: an integrated software package for processing of diffusion MRI data. 2010. ISMRM, 2010.
18. McAuliffe MJ, Lalonde FM, McGarry D, Gandler W, Csaky K, Trus BL. Medical Image Processing, Analysis and Visualization in clinical research. 2001:381-386; 26-27 July 2001. Proceedings 14th IEEE Symposium on Computer-Based Medical Systems. CBMS 2001. ISBN: 1063-7125. doi:10.1109/CBMS.2001.941749.
19. Özarslan E, Koay CG, Shepherd TM, et al. Mean apparent propagator (MAP) MRI: a novel diffusion imaging method for mapping tissue microstructure. *NeuroImage*. Sep 2013;78:16-32. doi:10.1016/j.neuroimage.2013.04.016
20. Pierpaoli C, Jones DK. Removing CSF Contamination in Brain DT-MRIs by Using a Two-Compartment Tensor Model. Proc. Intl. Soc. Mag. Reson. Med. 11. 2004:
21. Irfanoglu MO, Nayak A, Jenkins J, et al. DR-TAMAS: Diffeomorphic Registration for Tensor Accurate Alignment of Anatomical Structures. *NeuroImage*. May 15 2016;132:439-454. doi:10.1016/j.neuroimage.2016.02.066
22. Oishi K, Zilles K, Amunts K, et al. Human brain white matter atlas: identification and assignment of common anatomical structures in superficial white matter. *NeuroImage*. Nov 15 2008;43(3):447-57. doi:10.1016/j.neuroimage.2008.07.009
23. Nayak A, Walker L, Pierpaoli C, and The Brain Development Cooperative Group. Evaluation of pre-defined atlas based ROIs for the analysis of DTI data in Normal Brain Development. Proc. Intl. Soc. Mag. Reson. Med. 20. 2012:
24. Irfanoglu MO, Beyh A, Catani M, Dell'Acqua F, Pierpaoli C. ReImagining the Young Adult Human Connectome Project (HCP) Diffusion MRI Dataset. 2022:
25. Avants BB, Tustison NJ, Song G, Cook PA, Klein A, Gee JC. A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage*. Feb 1 2011;54(3):2033-44. doi:10.1016/j.neuroimage.2010.09.025
26. Yushkevich PA, Piven J, Hazlett HC, et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *NeuroImage*. Jul 1 2006;31(3):1116-28. doi:10.1016/j.neuroimage.2006.01.015
27. Srinivasan D, Erus G, Doshi J, et al. A comparison of Freesurfer and multi-atlas MUSE for brain anatomy segmentation: Findings about size and age bias, and inter-scanner stability in multi-site aging studies. *NeuroImage*. Dec 2020;223:117248. doi:10.1016/j.neuroimage.2020.117248
28. Anomalous Health Incidents and the Health Incident Response Task Force. (https://www.state.gov/anomalous-health-incidents-and-the-health-incident-response-task-force/). November 5, 2021.
29. Chen G, Xiao Y, Taylor PA, et al. Handling Multiplicity in Neuroimaging Through Bayesian Lenses with Multilevel Modeling. *Neuroinformatics*. Oct 2019;17(4):515-545. doi:10.1007/s12021-018-9409-6
30. Office of the Director of National Intelligence. *Complementary Efforts On Anomalous Health Incidents; Executive Summary of the Experts Panel: https://www.dni.gov/files/ODNI/documents/assessments/2022_02_01_AHI_Executive_Summary_FINAL_Redacted.pdf. February 1, 2022*.
31. Smith SM, Jenkinson M, Woolrich MW, et al. Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*. 2004;23 Suppl 1:S208-19. doi:10.1016/j.neuroimage.2004.07.051
32. Winkler AM, Ridgway GR, Webster MA, Smith SM, Nichols TE. Permutation inference for the general linear model. *Neuroimage*. May 15 2014;92:381-97. doi:10.1016/j.neuroimage.2014.01.060

33. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B (Methodological)*. 1995;57(1):289-300. doi:10.1016/s0166-4328(01)00297-2
34. Hafiz R, Irfanoglu MO, Nayak A, Pierpaoli C. "Pscore": A Novel Percentile-Based Metric to Accurately Assess Individual Deviations in Non-Gaussian Distributions of Quantitative MRI Metrics. *Journal of Magnetic Resonance Imaging*. 2024/01/30 2024;n/a(n/a)doi:https://doi.org/10.1002/jmri.29248
35. Hafiz R, Nayak A, Irfanoglu MO, Chan L, Pierpaoli C. Using 'P-scores': a novel percentile-based normalization method to accurately assess individual deviation in heavily skewed neuroimaging data. 2023 ISMRM & ISMRT Annual Meeting & Exhibition, Toronto, Canada, Program Abstract Number #3781, ISSN# 1545-4428 | Published date: 19 May, 2023.
36. Chen G, Taylor PA, Cox RW. Is the statistic value all we should care about in neuroimaging? *NeuroImage*. Feb 15 2017;147:952-959. doi:10.1016/j.neuroimage.2016.09.066
37. *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/*. 2022.
38. Gelman A, Hill J, Yajima M. Why We (Usually) Don't Have to Worry About Multiple Comparisons. *Journal of Research on Educational Effectiveness*. 2012/04/01 2012;5(2):189-211. doi:10.1080/19345747.2011.618213
39. Bürkner P-C. brms : An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*. 2017;
40. Lewandowski D, Kurowicka D, Joe H. Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis*. 2009/10/01/ 2009;100(9):1989-2001. doi:https://doi.org/10.1016/j.jmva.2009.04.008
41. Koenker R. *Quantile Regression*. Econometric Society Monographs. Cambridge University Press; 2005.
42. Koenker R, Bassett G. Regression Quantiles. *Econometrica*. 1978;46(1):33-50. doi:10.2307/1913643
43. Sherwood B, Zhou AX, Weintraub S, Wang L. Using quantile regression to create baseline norms for neuropsychological tests. *Alzheimers Dement (Amst)*. 2016;2:12-8. doi:10.1016/j.dadm.2015.11.005
44. Lê S, Josse J, Husson F. FactoMineR: An R Package for Multivariate Analysis. *Journal of Statistical Software*. 03/18 2008;25(1):1 - 18. doi:10.18637/jss.v025.i01
45. Ogawa S, Lee TM, Kay AR, Tank DW. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences*. 1990;87(24):9868-9872. doi:10.1073/pnas.87.24.9868
46. Kalcher K, Huf W, Boubela RN, et al. Fully exploratory network independent component analysis of the 1000 functional connectomes database. *Frontiers in human neuroscience*. 2012;6:301. doi:10.3389/fnhum.2012.00301
47. Meier TB, Wildenberg JC, Liu J, et al. Parallel ICA identifies sub-components of resting state networks that covary with behavioral indices. *Frontiers in human neuroscience*. 2012;6:281. doi:10.3389/fnhum.2012.00281
48. Cole DM, Smith SM, Beckmann CF. Advances and pitfalls in the analysis and interpretation of resting-state FMRI data. *Frontiers in systems neuroscience*. 2010;4:8. doi:10.3389/fnsys.2010.00008
49. Greicius MD, Krasnow B, Reiss AL, Menon V. Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*. Jan 7 2003;100(1):253-8. doi:10.1073/pnas.0135058100

50. Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America*. Jul 5 2005;102(27):9673-8. doi:10.1073/pnas.0504136102
51. Fox MD, Snyder AZ, Zacks JM, Raichle ME. Coherent spontaneous activity accounts for trial-to-trial variability in human evoked brain responses. *Nature neuroscience*. Jan 2006;9(1):23-5. doi:10.1038/nn1616
52. De Luca M, Beckmann CF, De Stefano N, Matthews PM, Smith SM. fMRI resting state networks define distinct modes of long-distance interactions in the human brain. *NeuroImage*. Feb 15 2006;29(4):1359-67. doi:10.1016/j.neuroimage.2005.08.035
53. Biswal B, Yetkin FZ, Haughton VM, Hyde JS. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic resonance in medicine*. Oct 1995;34(4):537-41. doi:10.1002/mrm.1910340409
54. Lowe MJ, Mock BJ, Sorenson JA. Functional connectivity in single and multislice echoplanar imaging using resting-state fluctuations. *NeuroImage*. Feb 1998;7(2):119-32. doi:10.1006/nimg.1997.0315
55. Cox RW. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and biomedical research, an international journal*. Jun 1996;29(3):162-73. doi:10.1006/cbmr.1996.0014
56. Cox RW, Hyde JS. Software tools for analysis and visualization of fMRI data. *NMR Biomed*. Jun-Aug 1997;10(4-5):171-8. doi:10.1002/(sici)1099-1492(199706/08)10:4/5<171::aid-nbm453>3.0.co;2-l
57. Davey CE, Grayden DB, Egan GF, Johnston LA. Filtering induces correlation in fMRI resting state data. *NeuroImage*. 2013/01/01/ 2013;64:728-740. doi:https://doi.org/10.1016/j.neuroimage.2012.08.022
58. Shirer WR, Jiang H, Price CM, Ng B, Greicius MD. Optimization of rs-fMRI Pre-processing for Enhanced Signal-Noise Separation, Test-Retest Reliability, and Group Discrimination. *NeuroImage*. 2015/08/15/ 2015;117:67-79. doi:https://doi.org/10.1016/j.neuroimage.2015.05.015
59. Jo HJ, Reynolds RC, Gotts SJ, et al. Fast detection and reduction of local transient artifacts in resting-state fMRI. *Computers in Biology and Medicine*. 2020/05/01/ 2020;120:103742. doi:https://doi.org/10.1016/j.compbiomed.2020.103742
60. Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*. 2012;59(3):2142-2154. doi:10.1016/j.neuroimage.2011.10.018
61. Reynolds RC, Taylor PA, Glen DR. Quality control practices in FMRI analysis: Philosophy, methods and examples using AFNI. *Frontiers in neuroscience*. 2022;16:1073800. doi:10.3389/fnins.2022.1073800
62. Shirer WR, Ryali S, Rykhlevskaia E, Menon V, Greicius MD. Decoding subject-driven cognitive states with whole-brain connectivity patterns. *Cerebral cortex (New York, NY : 1991)*. Jan 2012;22(1):158-65. doi:10.1093/cercor/bhr099
63. Beck AT, Epstein N, Brown G, Steer RA. An inventory for measuring clinical anxiety: psychometric properties. *J Consult Clin Psychol*. Dec 1988;56(6):893-7. doi:10.1037//0022-006x.56.6.893
64. Beck AT, Epstein N, Brown G, Steer R. Beck anxiety inventory. *Journal of consulting and clinical psychology*. 1993;
65. Weathers FW, Litz BT, Keane TM, Palmieri PA, Marx BP, Schnurr PP. The PTSD Checklist for DSM-5 (PCL-5). 2013; Retrieved from the National Center for PTSD website: http://www.ptsd.va.gov.

66. Blevins CA, Weathers FW, Davis MT, Witte TK, Domino JL. The Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5): Development and Initial Psychometric Evaluation. *Journal of Traumatic Stress*. 2015;28(6):489-498. doi:https://doi.org/10.1002/jts.22059
67. Glover GH. Overview of functional magnetic resonance imaging. *Neurosurgery clinics of North America*. Apr 2011;22(2):133-9, vii. doi:10.1016/j.nec.2010.11.001
68. Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America*. 2005;102(27):9673-9678. doi:10.1073/pnas.0504136102
69. Friston KJ. Functional and effective connectivity in neuroimaging: A synthesis. *Human brain mapping*. 1994;2(1-2):56-78. doi:https://doi.org/10.1002/hbm.460020107
70. Friston KJ, Frith CD, Liddle PF, Frackowiak RS. Functional connectivity: the principal-component analysis of large (PET) data sets. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism*. Jan 1993;13(1):5-14. doi:10.1038/jcbfm.1993.4
71. Abdi H, Williams LJ. Principal component analysis. *WIREs Computational Statistics*. 2010;2(4):433-459. doi:https://doi.org/10.1002/wics.101