Supplementary Materials: Brain Network Architecture Constrains Age-related Cortical Thinning

Marvin Petersen[1], Felix L. Nägele[1], Carola Mayer[1], Maximilian Schell[1], D. Leander Rimmele[1], Elina Petersen[2,3], Simone Kühn[4], Jürgen Gallinat[4], Uta Hanning[5], Jens Fiehler[5], Raphael Twerenbold[2,3,6,7], Christian Gerloff[1], Götz Thomalla[1], Bastian Cheng[1]

[1] Department of Neurology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

[2] Epidemiological Study Center, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

[3] Department of General and Interventional Cardiology, University Heart and Vascular Center, Hamburg, Germany

[4] Department of Psychiatry and Psychotherapy, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

[5] Department of Diagnostic and Interventional Neuroradiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

[6] German Center for Cardiovascular Research (DZHK), partner site Hamburg/Kiel/Luebeck, Hamburg, Germany

[7] University Center of Cardiovascular Science, University Heart and Vascular Center, Hamburg, Germany

Contents

S1: Functional processing stream of HCHS data

Results included in this manuscript come from preprocessing performed using *fMRIPrep* 20.2.6[1,2] , which is based on *Nipype* 1.7.0[3].

Anatomical data preprocessing
A total of 1 T1-weighted (T1w) images were found within the input BIDS dataset.The T1-weighted (T1w) image was corrected for intensity non-uniformity (INU) with N4BiasFieldCorrection[4], distributed with ANTs 2.3.3 (https://github.com/ANTsX/ANTs), and used as T1w-reference throughout the workflow. The T1w-reference was then skull-stripped with a *Nipype* implementation of the antsBrainExtraction.sh workflow (from ANTs), using OASIS30ANTs as target template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using fast[5]. Brain surfaces were reconstructed using recon-all[6] [FreeSurfer 6.0.1], and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical gray-matter of Mindboggle[7]. Volume-based spatial normalization to two standard spaces (MNI152NLin6Asym, MNI152NLin2009cAsym) was performed through nonlinear registration with antsRegistration (ANTs 2.3.3), using brain-extracted versions of both T1w reference and the T1w template. The following templates were selected for spatial normalization: *FSL's MNI ICBM 152 non-linear 6th Generation Asymmetric Average Brain Stereotaxic Registration Model* [RRID:SCR_002823; TemplateFlow ID: MNI152NLin6Asym], *ICBM 152 Nonlinear Asymmetrical template version 2009c* [RRID:SCR_008796; TemplateFlow ID: MNI152NLin2009cAsym],

Functional data preprocessing
For each of the 1 BOLD runs found per subject (across all tasks and sessions), the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. A deformation field to correct for susceptibility distortions was estimated based on *fMRIPrep*'s *fieldmap-less* approach. The deformation field results from co-registering the BOLD reference to the same-subject T1w-reference with its intensity inverted[8]. Registration was performed with antsRegistration (ANTs 2.3.3), and the process was regularized by constraining deformation to be nonzero only along the phase-encoding direction, and was modulated with an average fieldmap template. Based on the estimated susceptibility distortion, a corrected EPI (echo-planar imaging) reference was calculated for a more accurate co-registration with the anatomical reference. The BOLD reference was then co-registered to the T1w reference using bbregister (FreeSurfer) which implements boundary-based registration[9]. Co-registration was configured with six degrees of freedom. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) were estimated before any spatiotemporal filtering using mcflirt[10]. BOLD runs were slice-time corrected to 1.21s (0.5 of slice acquisition range 0s-2.42s) using 3dTshift from AFNI 20160207[11]. The BOLD time-series were resampled onto the following surfaces (FreeSurfer reconstruction nomenclature): *fsnative*, *fsaverage*, *fsaverage5*. The BOLD time-series (including slice-timing correction when applied) were resampled onto their original, native space by applying a single, composite transform to correct for head-motion and

susceptibility distortions. These resampled BOLD time-series will be referred to as *preprocessed BOLD in original space*, or just *preprocessed BOLD*. The BOLD time-series were resampled into several standard spaces, correspondingly generating the following *spatially-normalized, preprocessed BOLD runs*: MNI152NLin6Asym, MNI152NLin2009cAsym. First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. *Grayordinates* files[12] containing 91k samples were also generated using the highest-resolution fsaverage as intermediate standardized surface space. Automatic removal of motion artifacts using independent component analysis[13] was performed on the *preprocessed BOLD on MNI space* time-series after removal of non-steady state volumes and spatial smoothing with an isotropic, Gaussian kernel of 6mm FWHM (full-width half-maximum). Corresponding "non-aggressively" denoised runs were produced after such smoothing. Additionally, the "aggressive" noise-regressors were collected and placed in the corresponding confounds file. Several confounding time-series were calculated based on the *preprocessed BOLD*: framewise displacement (FD), DVARS and three region-wise global signals. FD was computed using two formulations following Power (absolute sum of relative motions[14]) and Jenkinson (relative root mean square displacement between affines[10]). FD and DVARS were calculated for each functional run, both using their implementations in *Nipype* [following the definitions by [14]]. The three global signals were extracted within the CSF, the WM, and the whole-brain masks. Additionally, a set of physiological regressors were extracted to allow for component-based noise correction [CompCor[15]]. Principal components were estimated after high-pass filtering the *preprocessed BOLD* time-series (using a discrete cosine filter with 128s cut-off) for the two *CompCor* variants: temporal (tCompCor) and anatomical (aCompCor). tCompCor components were then calculated from the top 2% variable voxels within the brain mask. For aCompCor, three probabilistic masks (CSF, WM and combined CSF+WM) were generated in anatomical space. The implementation differs from that of Behzadi et al. in that instead of eroding the masks by 2 pixels on BOLD space, the aCompCor masks were subtracted by a mask of pixels that contains a volume fraction of GM. This mask was obtained by dilating a GM mask extracted from the FreeSurfer's *aseg* segmentation, and it ensured that components were not extracted from voxels containing a minimal fraction of GM. Finally, these masks were resampled into BOLD space and binarized by thresholding at 0.99 (as in the original implementation). Components were also calculated separately within the WM and CSF masks. For each CompCor decomposition, the k components with the largest singular values were retained, so that the retained components' time series were sufficient to explain 50 percent of variance across the nuisance mask (CSF, WM, combined, or temporal). The remaining components were dropped from consideration. The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file. The confound time series derived from head motion estimates and global signals were expanded with the inclusion of temporal derivatives and quadratic terms for each[16]. Frames that exceeded a threshold of 0.5 mm FD or 1.5 standardised DVARS were annotated as motion outliers. All resamplings could be performed with *a single interpolation step* by composing all the pertinent transformations (i.e. head-motion transform matrices, susceptibility distortion correction when available, and co-registrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using antsApplyTransforms (ANTs), configured with Lanczos interpolation to

minimize the smoothing effects of other kernels. Non-gridded (surface) resamplings were performed using mri_vol2surf (FreeSurfer).

Many internal operations of *fMRIPrep* use *Nilearn*[17] 0.6.2, mostly within the functional processing workflow. For more details of the pipeline, see the section corresponding to workflows in *fMRIPrep*'s documentation.

Copyright waiver

S2: Structural processing stream of HCHS data

Preprocessing was performed using *QSIPrep* 0.14.2, which is based on *Nipype* 1.6.1[3].

Anatomical data preprocessing
The T1-weighted (T1w) image was corrected for intensity non-uniformity (INU) using N4BiasFieldCorrection[4] [ANTs 2.3.1], and used as T1w-reference throughout the workflow. The T1w-reference was then skull-stripped using antsBrainExtraction.sh (ANTs 2.3.1), using OASIS as target template. Spatial normalization to the ICBM 152 Nonlinear Asymmetrical template version 2009c was performed through nonlinear registration with antsRegistration [ANTs 2.3.1, https://github.com/ANTsX/ANTs], using brain-extracted versions of both T1w volume and template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using FAST[5] [FSL 6.0.3:b862cdd5].

Diffusion data preprocessing
Any images with a b-value less than 100 s/mm^2 were treated as a *b*=0 image. MP-PCA denoising as implemented in MRtrix3's dwidenoise[18] was applied with a 5-voxel window. After MP-PCA, Gibbs unringing was performed using MRtrix3's mrdegibbs[19]. Following unringing, B1 field inhomogeneity was corrected using dwibiascorrect from MRtrix3 with the N4 algorithm[4]. After B1 bias correction, the mean intensity of the DWI series was adjusted so all the mean intensity of the b=0 images matched across each separate DWI scanning sequence.

FSL (version 6.0.3:b862cdd5)'s eddy was used for head motion correction and Eddy current correction[20]. Eddy was configured with a *q*-space smoothing factor of 10, a total of 5 iterations, and 1000 voxels used to estimate hyperparameters. A linear first level model and a linear second level model were used to characterize Eddy current-related spatial distortion. *q*-space coordinates were forcefully assigned to shells. Field offset was attempted to be separated from subject movement. Shells were aligned post-eddy. Eddy's outlier replacement was run. Data were grouped by slice, only including values from slices determined to contain at least 250 intracerebral voxels. Groups deviating by more than 4 standard deviations from the prediction had their data replaced with imputed values. Final interpolation was performed using the jac method.

A deformation field to correct for susceptibility distortions was estimated based on *fmriprep*'s *fieldmap-less* approach. The deformation field is resulting from co-registering the b0 reference to the same-subject T1w-reference with its intensity inverted[8]. Registration was performed with antsRegistration (ANTs 2.3.1), and the process regularized by constraining deformation to be nonzero only along the phase-encoding direction, and modulated with an average fieldmap template. Based on the estimated susceptibility distortion, an unwarped b=0 reference was calculated for a more accurate co-registration with the anatomical reference. Several confounding time-series were calculated based on the preprocessed DWI: framewise displacement (FD) using the implementation in *Nipype* [following the definitions by [14]]. The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file. Slicewise cross correlation

was also calculated. The DWI time-series were resampled to ACPC, generating a *preprocessed DWI run in ACPC space* with 2mm isotropic voxels.

MRtrix3 Reconstruction

Multi-tissue fiber response functions were estimated using the dhollander algorithm. FODs were estimated via constrained spherical deconvolution[21] (CSD) using an unsupervised multi-tissue method[22,23]. A single-shell-optimized multi-tissue CSD was performed using MRtrix3Tissue (https://3Tissue.github.io), a fork of MRtrix3[24] FODs were intensity-normalized using mtnormalize[25].

Many internal operations of *QSIPrep* use *Nilearn*[17] and *Dipy*[26]. For more details of the pipeline, see the section corresponding to workflows in *QSIPrep*'s documentation.

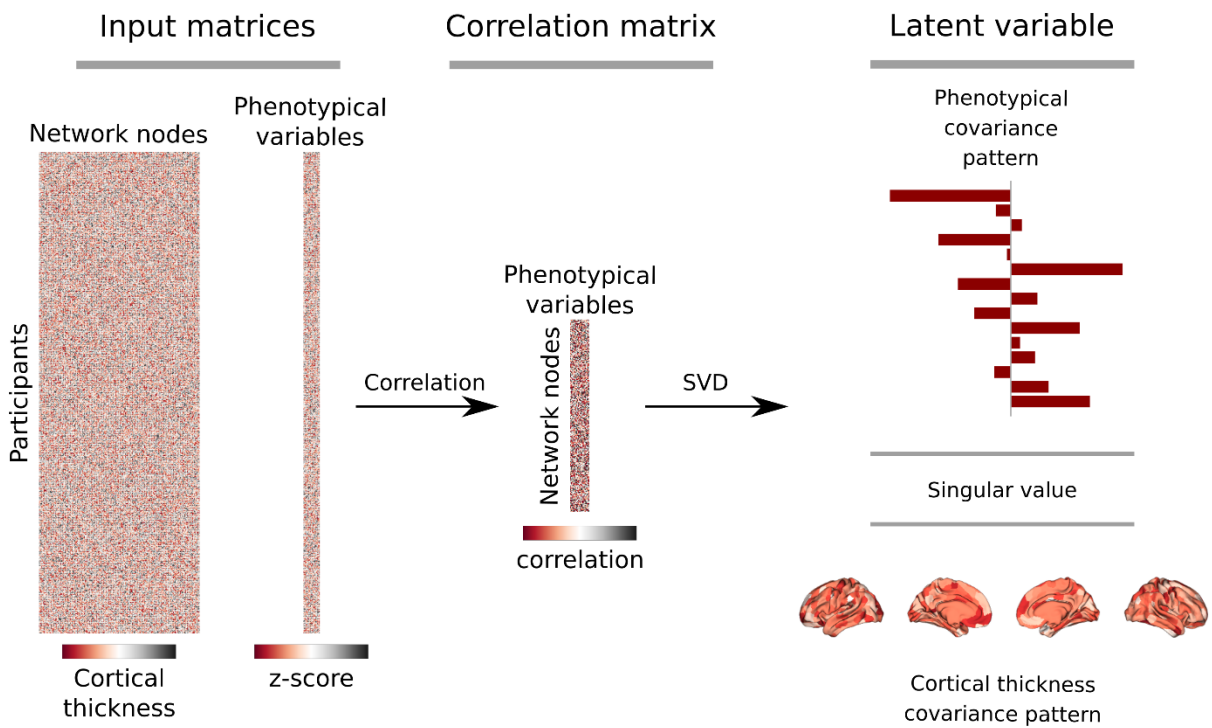## S3: Illustration partial least squares analysis

Figure S3: Partial least squares methodology. Please refer to S4 for a detailed description of the approach. This figure is a modification from figure 1 of Zeighami et al., 2019[27].

S4: Detailed explanation partial least squares analysis

Thickness and phenotypical data were arranged in two matrices $X_{n_{participants} \times n_{nodes}}$ and $Y_{n_{participants} \times n_{phenotypical\ variables}}$, z-scored and the thickness-phenotypical correlation matrix was then calculated which encompassed the correlation of all nodal cortical thickness values and phenotypical measures across subjects. Upon that singular value decomposition was performed on the correlation matrix resulting in a set of mutually orthogonal latent variables (LVs). The number of latent variables is determined by the smaller dimension of the covariance matrix - i.e., its rank; here the number of phenotypical variables. A latent variable consists of a left singular vector, a right singular vector and a singular value. Singular value decomposition results in matrices of left ($U_{n_{nodes} \times n_{phenotypical\ variables}}$) and right ($V_{n_{phenotypical\ variables} \times n_{phenotypical\ variables}}$) singular vectors as well as a diagonal matrix of singular values ($\Delta_{n_{phenotypical\ variables} \times n_{phenotypical\ variables}}$). A singular vector weights the corresponding original variables so that their covariance is maximized. The weighted nodal values can be understood as maximally covarying thickness patterns and their corresponding manifestation phenotypical data. Each pairing of left and right singular vector is connected by a singular value which is propotional to the covariance of both. The explained covariance of a latent variable was assessed as the ratio of the squared singular value to the sum of the other squared singular values. Significance of a latent variable was assessed by comparing the observed explained variance to a non-parametric distribution of permuted values via permutation (n=5000) of the subject-order in X. To measure to which extent a covariance pattern represented by a latent variable is manifested within a subject, participant-specific scores were calculated by projecting U on X for a thickness score

$$Thickness\ score = UX$$

and V on Y for a phenotypical score

$$Phenotypical\ score = VY$$

which can be thought of as a component or factor weighting in principal component analysis or factor analysis, respectively. Bootstrap resampling was used to identify nodes with a considerable contribution to the thickness-phenotypical relationship. For that, participants were randomly sampled with replacement from X and Y (n=5000) which results in a set of resampled covariance matrices propagated to singular value decomposition. This yields a sampling distribution for each node enabling the computation of a bootstrap ratio as the ratio of its singular vector weight and the corresponding standard error as estimated from bootstrapping. The bootstrap ratio is used to measure a node's contribution to the observed covariance pattern represented by a latent variable as relevant nodes exhibit a high singular vector weight and have a small standard error which means they are stable across subjects. The bootstrap ratio pattern was spatially correlated with the thickness β-map to assess whether the age-related cortical thinning pattern corresponds with the nodes contributing to the covariance pattern.

S5: Demographics of age ranges

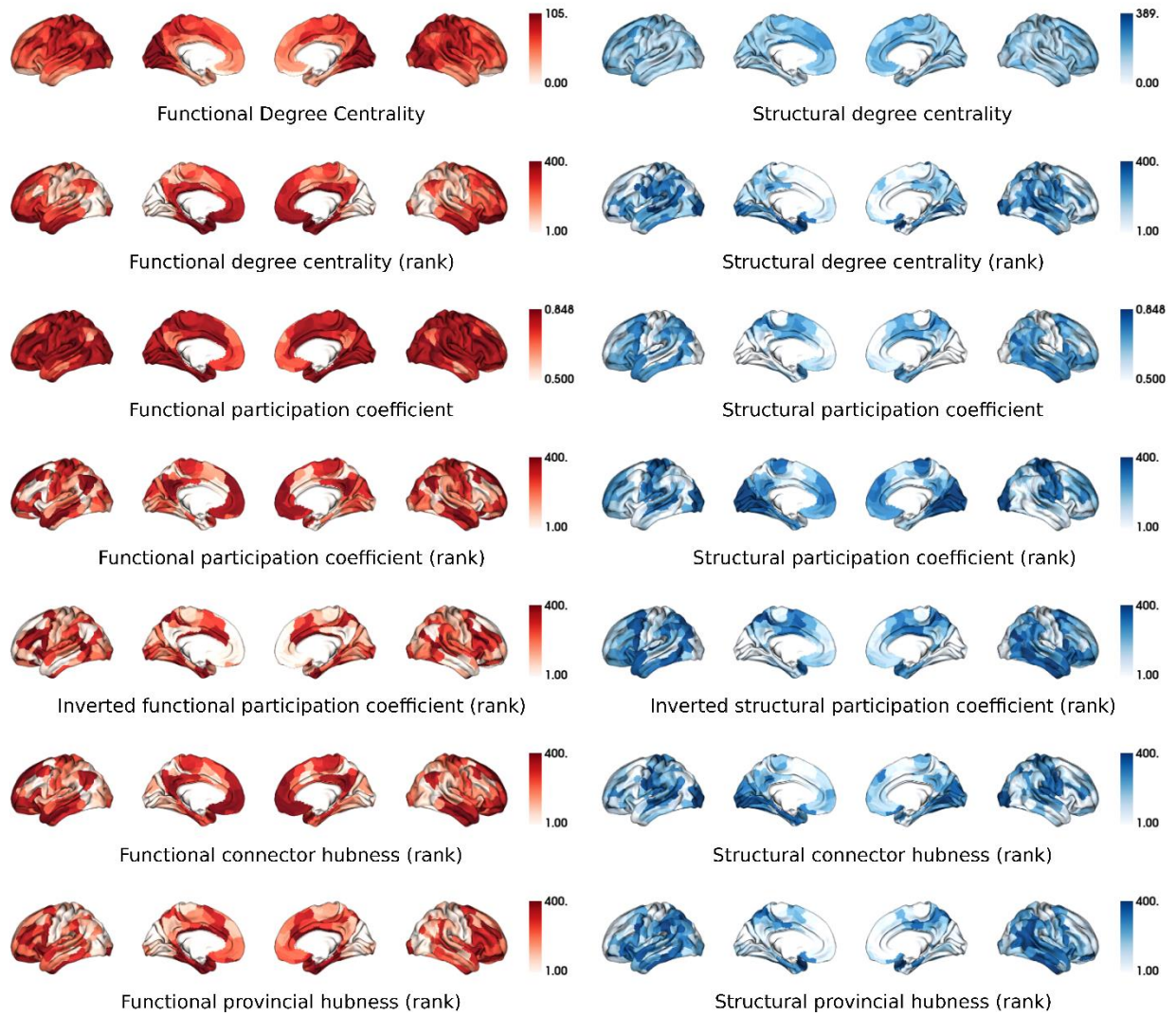|	45-50	50-55	55-60	60-65	65-70	70-75	75-80
n	150	286	366	407	543	602	220
Age, median (IQR)	48.00 (2)	52.00 (2)	57.00 (2)	62.00 (2)	67.00 (2)	72.00 (2)	76.00 (2)
Sex, % female	46.00%	50.00%	46.00%	41.00%	39.00%	43.00%	46.00%
Years of education, median (IQR)	14.00 (4)	15.00 (4)	14.00 (5)	13.50 (5)	13.00 (4)	13.00 (3)	13.00 (3)

S6: Hubness measures

Figure S6: Surface plots displaying all metrics relevant for the hubness analysis.

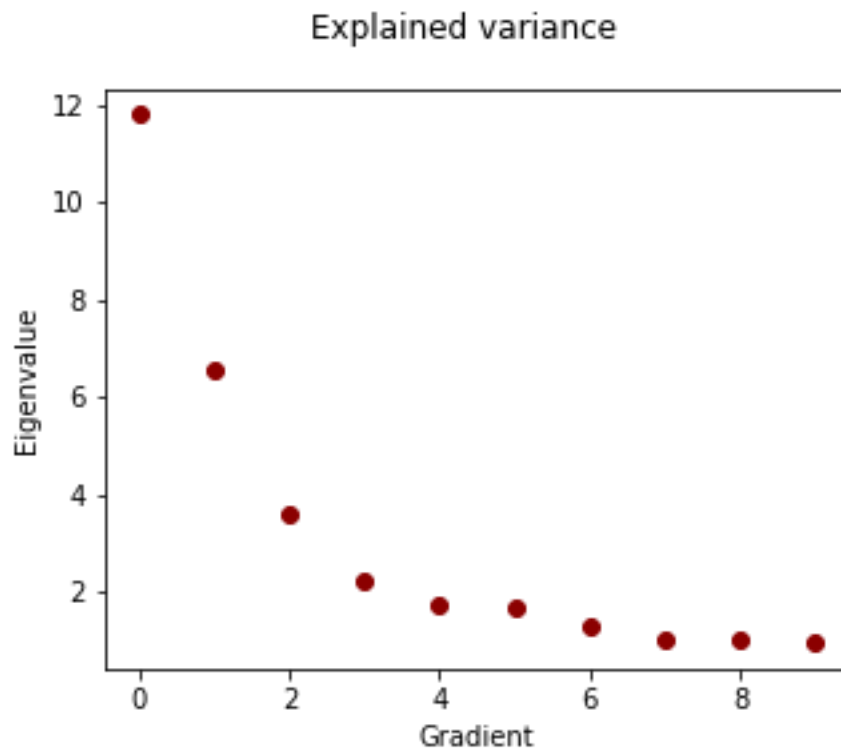**S7: Scree plot for explained variance of functional gradients**

Figure S7: Scree plot of respective eigenvalues of derived functional gradients. High eigenvalues denote high explained variance.

S8: Sensitivity analysis: functional and structural 2nd step neighborhood alteration
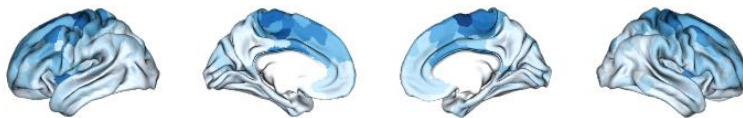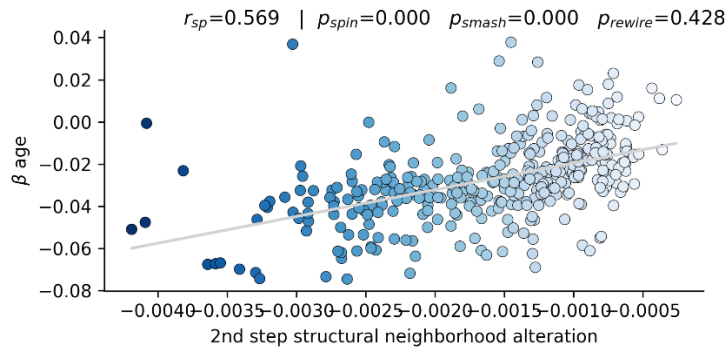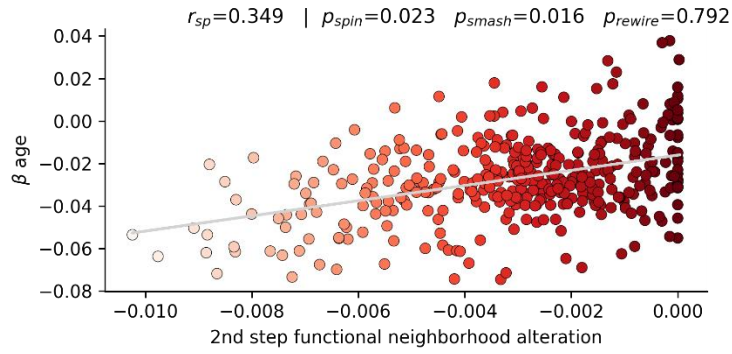

Figure S8: Scatter plots as well as surface visualizations of the functional (red) and structural (blue) 2nd step neighborhood alteration.

S9: Sensitivity analysis: different connectomes and Schaefer parcellations

	Functional Connector Hubness	Structural Connector Hubness	Functional Provincial Hubness	Structural Provincial Hubness	Functionally-defined Neighborh.	Structurally-defined Neighborh.	Functional Gradient 1	Functional Gradient 2	Structure-function Coupling
r_{sp} Schaefer400 HCHS	-0.07	-0.007	-0.02	0.26	0.41	0.44	0.25	0.47	-0.25
P_{spin} Schaefer400 HCHS	0.294	0.431	0.46	0	0.007	0	0.042	0	0.069
r_{sp} Schaefer100 HCP	-0.054	0.221	0.11	0.41	0.13	0.45	-0.33	0.05	-0.1
P_{spin} Schaefer100 HCP	0.370	0.062	0.308	0.004	0.292	0.033	0.049	0.408	0.177
r_{sp} Schaefer200 HCP	-0.152	0.017	0.22	0.24	0.16	0.48	-0.32	0.46	-0.17
P_{spin} Schaefer200 HCP	0.100	0.432	0.115	0.037	0.225	0.003	0.01	0	0.085
r_{sp} Schaefer400 HCP	-0.119	0.029	0.27	0.22	0.23	0.64	-0.34	0.44	-0.16
p_{spin} Schaefer400 HCP	0.2	0.395	0.029	0.047	0.118	0	0.005	0	0.123

S10: Partial least squares analysis: latent variable statistics

Latent variable	Explained variance (%)	p-value
0	85.096707	0.000200
1	9.171140	0.000400
2	1.516065	0.039992
3	0.832268	0.188562
4	0.689248	0.729054
5	0.620280	0.434513
6	0.521711	0.733253
7	0.443405	0.677465
8	0.361422	0.926415
9	0.297952	0.972006
10	0.199595	0.668666
11	0.191003	0.999800
12	0.059203	0.956809

S11: Partial least squares analysis: remaining significant latent variables of the partial least squares analysis
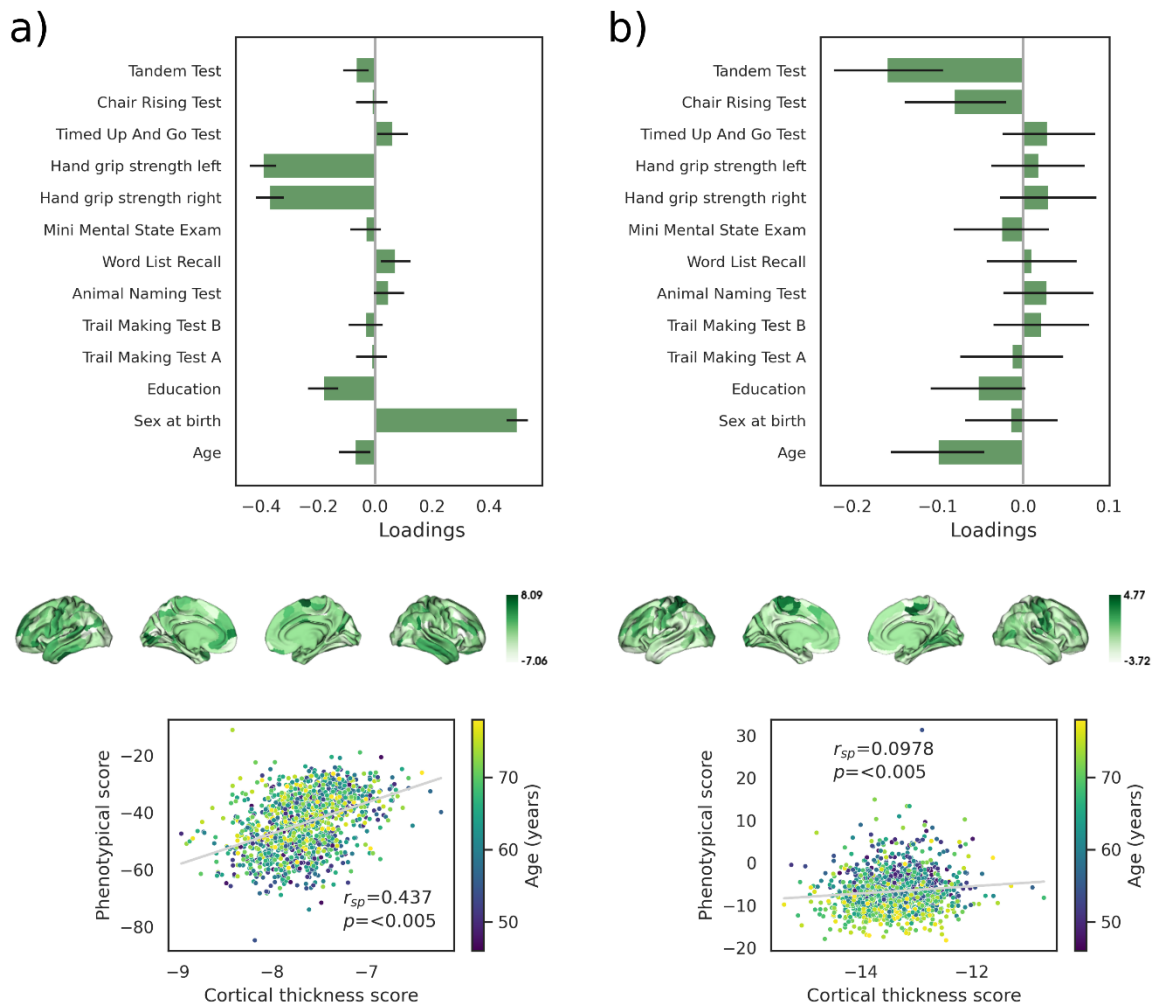

Figure S11. Partial least squares analysis results of the significant latent variables 2 (a) and 3 (b)

S12: Partial least squares analysis: Cross-validation results

Out-of-sample correlation between cortical thickness and phenotypical scores

CV folds	Spearman r	FDR-corrected p
0	0.332502	0.000166
1	0.334353	0.000166
2	0.392078	1.44E-05
3	0.423133	2.93E-06
4	0.143506	0.106086
5	0.33646	0.000166
6	0.331384	0.000166
7	0.379984	2.42E-05
8	0.538506	5.51E-10
9	0.299848	0.00065

S13: Partial least squares analysis without age, sex and education
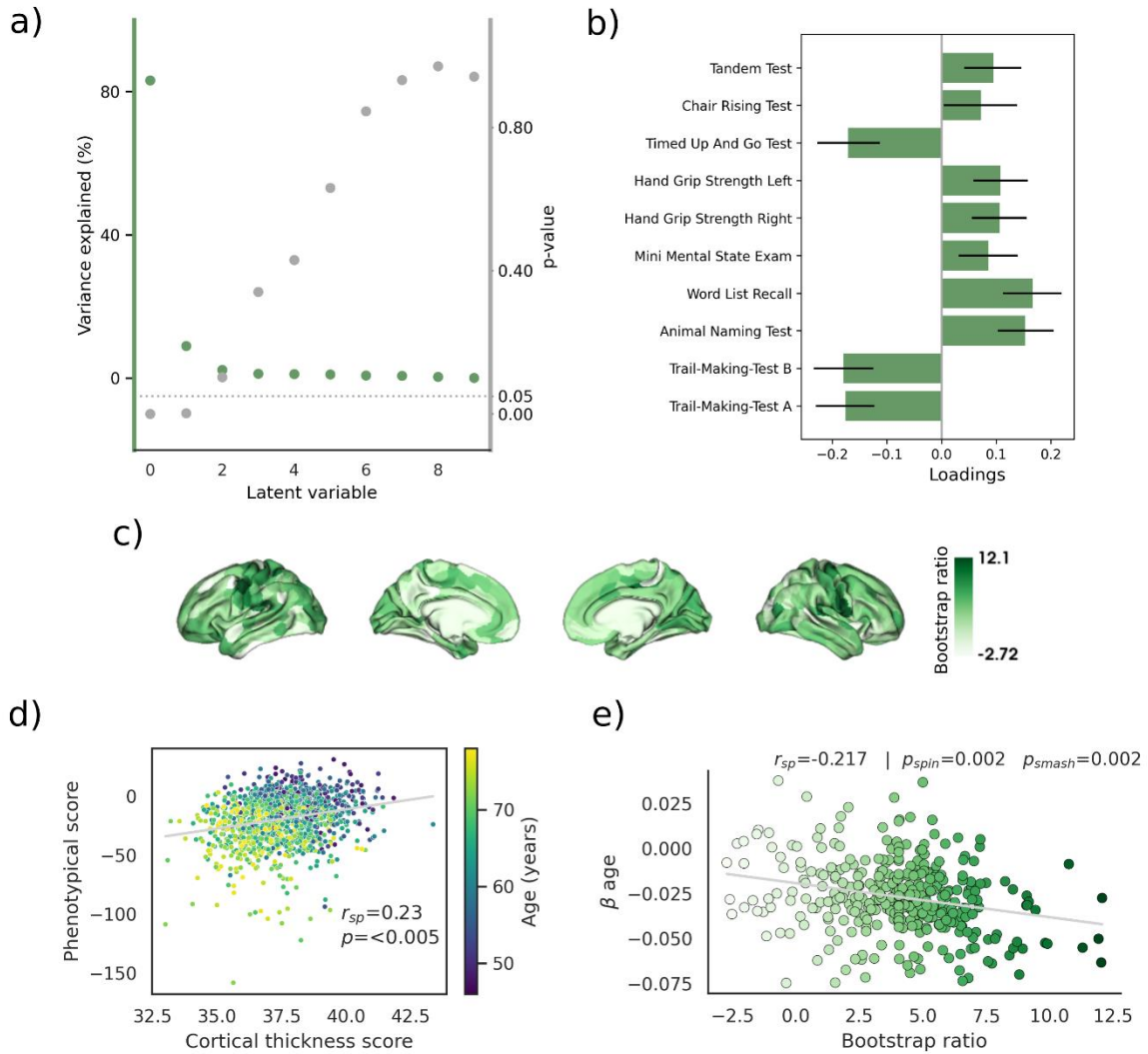

Figure S13. Partial least squares analysis results without consideration of age, sex and education as phenotypical variables. Subplot order matches that in *figure 5* of the main text.

S14: Partial least squares analysis after deconfounding cognitive and motor scores for age
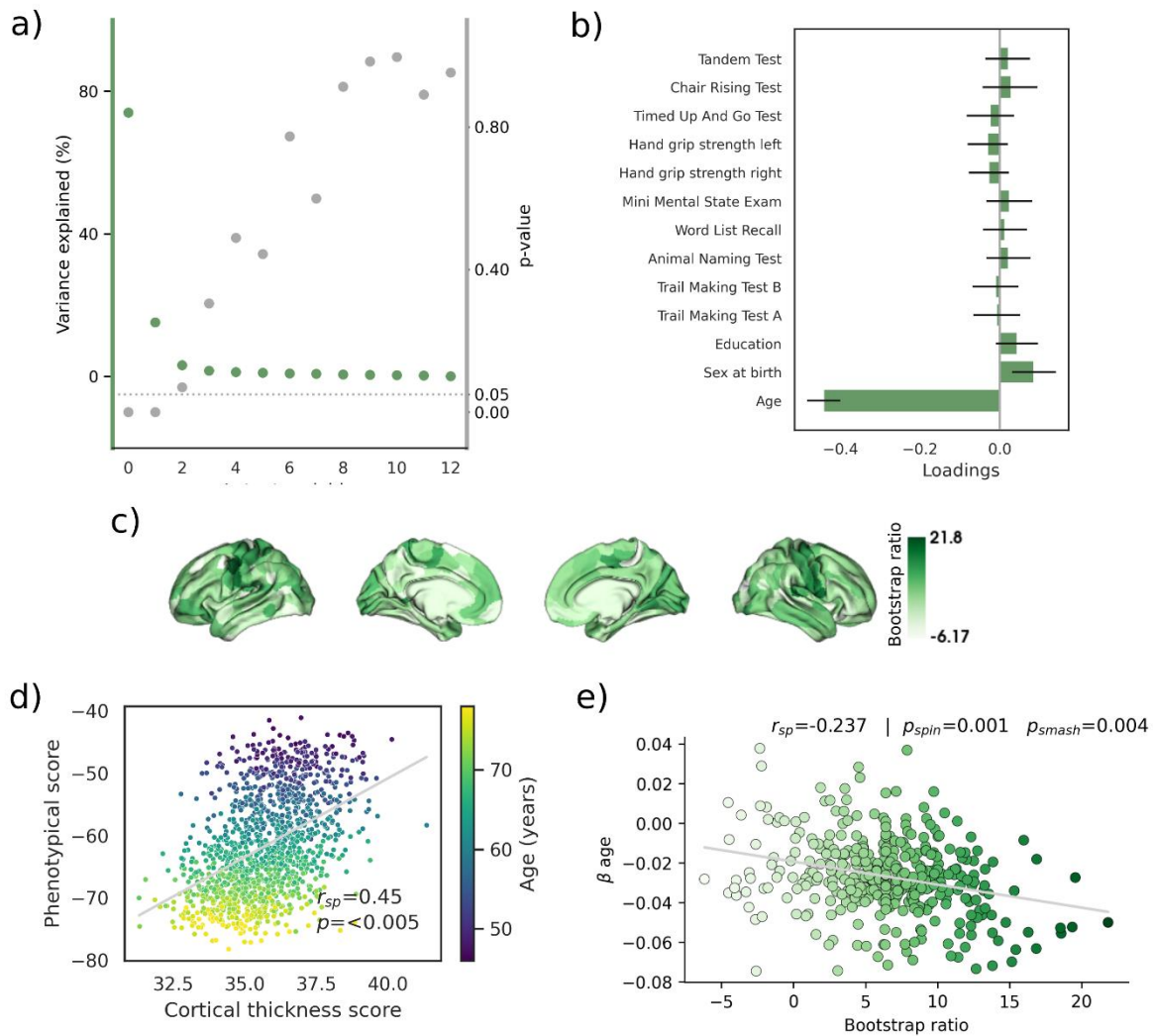

Figure S14. Partial least squares analysis results after deconfounding cognitive and motor scores for age. Subplot order matches that in *figure 5* of the main text.

S15: Code availability

Analysis step	Code URL
DWI preprocessing with QSIPrep	https://github.com/csi-hamburg/CSIframe/blob/709275c816b7746bf7168f69b652b2aec569b838/pipelines/qsiprep/qsiprep.sh
fMRI preprocessing with fMRIPrep	https://github.com/csi-hamburg/CSIframe/blob/main/pipelines/fmriprep/fmriprep.sh
Functional connectome reconstruction	https://github.com/csi-hamburg/CSIframe/blob/main/pipelines/xcpengine/xcpengine.sh
Spatial correlations, partial least squares analysis	https://github.com/csi-hamburg/2022_petersen_aging_ct_connectivity
Structural processing with CAT	https://github.com/csi-hamburg/CSIframe/blob/709275c816b7746bf7168f69b652b2aec569b838/pipelines/cat12/cat12.sh

References

1. Esteban, O. *et al.* fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nature Methods* **16**, 111–116 (2019).

2. Esteban, O. *et al.* Analysis of task-based functional MRI data preprocessed with fMRIPrep. *Nat Protoc* **15**, 2186–2202 (2020).

3. Gorgolewski, K. *et al.* Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python. *Frontiers in Neuroinformatics* **5**, (2011).

4. Tustison, N. J. *et al.* N4ITK: improved N3 bias correction. *IEEE Trans Med Imaging* **29**, 1310–1320 (2010).

5. Zhang, Y., Brady, M. & Smith, S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans Med Imaging* **20**, 45–57 (2001).

6. Dale, A. M., Fischl, B. & Sereno, M. I. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage* **9**, 179–194 (1999).

7. Klein, A. *et al.* Mindboggling morphometry of human brains. **13**, (2017).

8. Huntenburg, J. M., Gorgolewski, K. J., Anwander, A. & Margulies, D. Evaluating nonlinear coregistration of BOLD EPI and T1 images. (2014) doi:10.7490/F1000RESEARCH.1096036.1.

9. Greve, D. N. & Fischl, B. Accurate and robust brain image alignment using boundary-based registration. *NeuroImage* **48**, 63–72 (2009).

10. Jenkinson, M., Bannister, P., Brady, M. & Smith, S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* **17**, 825–841 (2002).

11. Cox, R. W. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res* **29**, 162–173 (1996).

12. Glasser, M. F. *et al.* The Human Connectome Project's neuroimaging approach. *Nat Neurosci* **19**, 1175–1187 (2016).

13. Pruim, R. H. R., Mennes, M., Buitelaar, J. K. & Beckmann, C. F. Evaluation of ICA-AROMA and alternative strategies for motion artifact removal in resting state fMRI. *NeuroImage* **112**, 278–287 (2015).

14. Power, J. D. *et al.* Methods to detect, characterize, and remove motion artifact in resting state fMRI. *Neuroimage* **84**, 320–341 (2014).

15. Muschelli, J. *et al.* Reduction of motion-related artifacts in resting state fMRI using aCompCor. *Neuroimage* **96**, 22–35 (2014).

16. Satterthwaite, T. D. *et al.* Heterogeneous impact of motion on fundamental patterns of developmental changes in functional connectivity during youth. *Neuroimage* **83**, 45–57 (2013).

17. Abraham, A. *et al.* Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics* **8**, (2014).

18. Veraart, J. *et al.* Denoising of diffusion MRI using random matrix theory. *Neuroimage* **142**, 394–406 (2016).

19. Kellner, E., Dhital, B., Kiselev, V. G. & Reisert, M. Gibbs-ringing artifact removal based on local subvoxel-shifts. *Magn Reson Med* **76**, 1574–1581 (2016).

20. Andersson, J. L. R. & Sotiropoulos, S. N. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *Neuroimage* **125**, 1063–1078 (2016).

21. Tournier, J.-D. *et al.* Resolving crossing fibres using constrained spherical deconvolution: validation using diffusion-weighted imaging phantom data. *Neuroimage* **42**, 617–625 (2008).

22. Dhollander, T. & Connelly, A. A novel iterative approach to reap the benefits of multi-tissue CSD from just single-shell (+b=0) diffusion MRI data. 8 (2016).

23. Dhollander, T., Mito, R., Raffelt, D. & Connelly, A. Improved white matter response function estimation for 3-tissue constrained spherical deconvolution. 10 (2019).

24. Tournier, J.-D. *et al.* MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. *NeuroImage* **202**, 116137 (2019).

25. (ISMRM 2017) Bias Field Correction and Intensity Normalisation for Quantitative Analysis of Apparent Fibre Density. https://ismrm.gitlab.io/2017/3541.html.

26. Garyfallidis, E. *et al.* Dipy, a library for the analysis of diffusion MRI data. *Frontiers in Neuroinformatics* **8**, (2014).

27. Zeighami, Y. *et al.* A clinical-anatomical signature of Parkinson's disease identified with partial least squares and magnetic resonance imaging. *NeuroImage* **190**, 69–78 (2019).